

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 12:03:42 ; Search time 34846 Seconds  
(without alignments)  
11772.891 Million cell updates/sec

Title: PCT-US04-13989-18

Perfect score: 8675

Sequence: 1 aaccggcgtaactggsggtg.....ctggcaccgcgcatgtctct 8675

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank1:\*  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_rc:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2797.6	32.2	6815	1	AB076662 Pantoea a
2	2688.4	31.0	6586	1	AY166713 Pantoea s
3	2618	30.2	6918	1	D90087 Erwinia ure
4	2616.4	30.2	6918	6	I12910 Sequence 13
5	2486.2	28.7	6965	1	ERMCRTA M90698 Pantoea ags
6	2248	25.9	12753	1	ERMCRTS M87280 Pantoea ags
7	1885.6	21.7	4624	6	CQ793208 Sequence
8	1345.8	15.5	349107	1	BX571873 Photorhab
9	1265.8	14.6	2415	1	ERMCRITB M38423 Pantoea ags
10	1156.4	13.3	3801	1	AF408848 Xanthobac
11	1085.8	12.5	6635	1	AF218415 Bradyrhiz
12	1054.6	12.2	6335	1	PMCRWTYI Y15112 Paracoccus
13	1053.6	12.1	5373	1	ATUCRTWA D58420 Paracoccus
14	1032.2	11.9	5188	1	FAU62808 U62808 Flavobacter
15	1032.2	11.9	8625	6	AR169831 Sequence
16	1032.2	11.9	8625	6	AR452005 Sequence
17	1032.2	11.9	11233	6	AR169852 Sequence
18	1032.2	11.9	11233	6	AR452026 Sequence
19	826.2	9.5	1479	6	CQ793426 Sequence

20	826.2	9.5	1479	6	CQ793595	CQ793595 Sequence
21	826.2	9.5	1479	6	E22381	E22381 beta-Carote
22	826.2	9.5	1479	6	I12907	I12907 Sequence 10
23	826.2	9.5	1493	6	BD241820	BD241820 Method of
24	826.2	9.5	1493	6	AX014697	AX014697 Sequence
25	811.8	9.4	1479	6	AX394988	AX394988 Sequence
26	751.4	8.7	1518	6	I72653	I72653 Sequence 7
27	746.4	8.6	1522	6	I72654	I72654 Sequence 9
28	655.2	7.6	349723	1	BX842650	BX842650 Bellovib
29	611.4	7.0	1482	6	A58565	A58565 Sequence 6
30	611.4	7.0	1482	6	A84702	A84702 Sequence 12
31	611.4	7.0	1482	6	AR103081	AR103081 Sequence
32	611.4	7.0	1482	6	AR139697	AR139697 Sequence
33	611.4	7.0	1482	6	AR391876	AR391876 Sequence
34	562	6.5	1198	6	I24838	I24838 Sequence 5
35	562	6.5	1198	6	I72652	I72652 Sequence 5
36	504.4	5.8	1083	6	I40224	I40224 Sequence 1
37	483	5.6	349737	1	BX572597	BX572597 Rhodosp
38	466.2	5.4	30040	1	AF182374	AF182374 Bradyrhiz
39	456.6	5.3	1296	6	AX394984	AX394984 Sequence
40	452.2	5.2	1235	6	I72656	I72656 Sequence 13
41	449.8	5.2	5397	12	AY605097	AY605097 Synthetic
42	441	5.1	1235	6	I72657	I72657 Sequence 15
43	440	5.1	1149	6	AX394986	AX394986 Sequence
44	429	4.9	1581	1	D83514	D83514 Erythroba
45	429	4.9	1581	6	E11103	E11103 Gene coding

## ALIGNMENTS

RESULT 1	AB076662	6815 bp	DNA	linear	BCT 26-DEC-2001
LOCUS	AB076662	Pantoea agglomerans pv. milletiae	crx, crtx, crty, crti, crtb,		
DEFINITION	AB076662	crx genes, complete cde.			
ACCESSION	AB076662.1	GI:18143444			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					

```
gene
1081..2376
/gene="crlx"
1081..2376
/gene="crlx"
/codon_start=1
/transl_table=11
/protein_id="BAB79601.1"
/db_xref="GI:18143446"
/translation="MSHFAVIAPPPYSHVOALQHLSQLIARGHQLFIQOTVSALT
TDSRIGFPGSLAPAGSLAHTLQLAHPLGSMKLINEMARSTDMLCRELPIVLS
TADIVGVNDOMBAGALAAALNLPVSVACALPIAREADPLAACPPEYATDQAR
ERTTSBKTYDMLRRHRDVIYARNAHMGALPREKHGCSFVAQINQLPEIDPEQO
TLPDHFVSGPLRTTEPAAPASOPRYEPHGDPRIPASLGLTQGHRYGLKFTIVACQ
EIDQHLALHAGRLSPQAEKLAQASHVQVDPADQAAALQADLVITGGNMVTDG
INHLPLTITPLAFDQPGVAAVVMHIGIRASRFTTSHMAQLCTLADESVQOM
KTRSMALROAGGTTLADIIVEQMLTRQPVLTRHYAAV"
2363..3523
/gene="crlx"
2363..3523
/gene="crlx"
/codon_start=1
/transl_table=11
/protein_id="BAB79602.1"
/db_xref="GI:18143447"
/translation="MPYRDLILVAGLNGULALRLKQRPRLILLIDAREBGANH
TWSFBADELTEQHRWIAPLVHHMPCYEVRFPQSRSMNGYFCVTAERFAQVIRP
FADPLINTRVASIASRAVTLDDGRVLESADVIDRGYCPDGLRMGFQSFQEOOL
SEBGLTAPILIMATVDDQAGYRFVSLPSADPLITEDHYIDNLTLEDGRARONIH
DYAOCQWRILDRLLERQALPITLGDVAAPFQNRDLPSCGACGIFHTTQYSLPL
AYALADRLAOMQITSETLTIATIQPRASQMOORFPRMIRMLFLAGRADQRYVQO
RFTYGLPREGLIARFYAGKTLPRRLILSGRPVPLVALQALMTPHRQAMO"
3520..4998
/gene="crl1"
3520..4998
/gene="crl1"
/codon_start=1
/transl_table=11
/protein_id="BAB79603.1"
/db_xref="GI:18143448"
/translation="MNRITVIGAFGGLALALRLQASGVPTRLLEORDKFGRAVYU
DOGPTPAGPTVITDPSAIRELFTIAGKSLDYVELMVPVRYRILCMESGKVSND
QALYEAQIAPNPRDVBGRRPFLAISAIVABGTYLKGCTPISFPRMIAAQLAKL
QARSVYSKVASYIEDHEKQATSFSLVGNPFAVSSLYTLIHLEBWGVMPFG
GTGALVOGWKTLFEDLGEVELNAAVRLTQENRITVAHLKQGRVETPAVANSNDY
VHTYRELISQHPASQAQGRSLQNKMSNSLFVIYFGLHHNOLAHTVFCGRYSEL
IDEIFNDGLAEDFSLYLAHPCVTDPSLAPEGGSYYLAVPVHLGADIDMAVEGR
LRRIPIYEQHYMPGLRSOLVTHRIPTPPRDELNAVQGSASFVEPIITQSAWFRP
HNDDKINNTYLVGAGTHPGAGIPGVIGSAKATAGMLBDA"
5034..5924
/gene="crlB"
5034..5924
/gene="crlB"
/codon_start=1
/transl_table=11
/protein_id="BAB79604.1"
/db_xref="GI:18143449"
/translation="MEVGSKFATASKLFGAKTRRSYMLYANCRHCDVDIVDQVLGF
SNDPISQSAEORLAQLEMKTROAVASQOMBPAFAFCVAAHNDILPYAADHLAG
FAMDVHETRYOTLDDTLRCYHAGVAGVMAAOIMGRDAVTLDRACDGLAPADINI
ARDIVEDADAGRCITPATWLAEBGLTRENADONQALSRVARLVEYTERPYRNAS
AGLPLGRSAMAIATIQOYVYKIGMKVQVQASOAMOROSTSTPEKALLVAASGA
VTSRVNARHARSADLWQRPV"
complement(5863..6390)
/gene="crlZ"
complement(5863..6390)
/gene="crlZ"
/codon_start=1
/transl_table=11
/protein_id="BAB79605.1"
/db_xref="GI:18143450"
/translation="MLMINVALIVLVIVIGMEITALLAHNYIHWGQWGNHLSHNEBH
```

## ORIGIN

Query Match 32.2%; Score 2797.6; DB 1; Length 6815;  
Best Local Similarity 66.5%; Pred. No. 0;  
Matches 4140; Conservative 0; Mismatches 2054; Indels 30; Gaps 8;

KCMFEVNDLXAVVPAALSTILYLGSTGVPLQWIGAKMTYGLLYEIVHDELHQRW  
PFRIVRRGLRLYIAHHRHNAVRKSGVSGFLYAPPLSKDQTLNERHVKXGA  
ARDQSRVERDAPPGK"

```
QY 1243 AAGCGATCTCTTCAGTTGCAATGAGAGCCCATCTTGAACATTATTCCTG 1302
DB 200 AAGCGATGAGCAGAGCTGTGAAGCAATGAGCAAGCGCTTGATACCTTTACCG 259
QY 1303 CCGGCGACAGAGCGCATCGCTTCGCGGCCCATCTGCTCGCGCAGCTTGGCA 1362
DB 260 TTGAAAGCGACGCTTATAGTGGGGCGTCCATGGCCGACGCTGGCGCACGGAA 319
QY 1363 AAGCTATTGCTCGCTCTGCTGCTGCGACAGCCGATATAGGCTGTGAGCGCG 1422
DB 320 AGCTTATCGTCCACTGCTGTCTTGTGCTGAGCGCGCCATCTGGCTGCAAGCCACGC 379
QY 1423 AGCAGGCAATCTTGAATCTTCTGCTGCGGTGCAAAATGTCAGCGTCTCATATCC 1482
DB 380 CTGCGCGCTGCTGATCTGCTGCGCGGTAGATGGTGCATTCGCGCATCTGATTC 439
QY 1483 TCGACGACATTCATGATGATGATTAAGCCCGGATGCGAGCTGGGCGCCGCGCATCCACT 1542
DB 440 TGGATGACATGACCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 499
QY 1543 GTGATATGAGGGAAGAAAGTGGCCGATCTGCGACGCGTGGCGCTACTCAGCGCGCTTTG 1602
DB 500 GCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559
QY 1603 AGGTGATGCTTCCGCGCGCGCGGTCTGCGAAGCAAGCAAAAGCCGATTCGCGAGC 1662
DB 560 GCGTGGTGGTGGCGGAGAAAGGCTTAAAGCGCAACCGCCAGAGCGCATCTGCGAGAT 619
QY 1663 TCTCTCTGCGCGTGGCGCTGCGAGGAGCTGTTAGGGTCACTTCCAGATCTGATGACG 1722
DB 620 TATCCACGAGTGGCGATGACAGAGGCTGTTGAGGGCAATTTAAGATCTTCCGAGG 679
QY 1723 GCGCACACAGCGCGCAGTCCGGAAGCCATCACCTGACCAATGATGAAACAGAGCTTC 1782
DB 680 GTACCAAGCCAGCAGCAGCGCTGACGCGCATCTGATGACCAATCTATMAAACAGCAGCG 739
QY 1783 TGTTCGCGCGCAGCTGACAGATGCGCGCATTTGCGGCGCATGCTGACAGTACGTC 1842
DB 740 TGTTCGCGCGCTCATGACAGATGCGCTCTATCTGCTGTAAGCGCTGAGTGAAGCGCG 799
QY 1843 AGCTTTAAGCTATTTTGGCAGAGATTTAGTACAGCTTTCCAGTTAATGAGACGCTGG 1902
DB 800 AACAGCTGACCGCTTTTGTCTTAACTTGTGTAGGCTTTCCAGCTTGAAGCATCTCA 859
QY 1903 CGGATGCTCTTAAGCAGACACCGGCAAGGACCTGTATGAGATCTAGGCAATTCACGCTGG 1962
DB 860 CTGACGCGCAGTGGCGGACACCGCGTAAAGATGCCCATGAGATGACGGGAATTCACGCTGG 919
QY 1963 TGCAGATGCTGGGCGCGGAAAGGGCTGAGCGTCTGCGCGCACTTAAGACGCGCG 2022
DB 920 TGAATCTGCTGGGACACAGCGGCTTGAACCGGACCTGCGCGCATCTGCGCTGGCGCA 979
QY 2023 ATGCAACCTGCGCTGCGCTGCGCATCGCGTCTGCGCGCACTTAATGATGACAGCGCC 2082
DB 980 GCGAGGATCTGTTATGCGCTGCGCAGAGCGTATGCCACACCACTTTGTTCAGGCGCT 1039
QY 2083 TGTTTATGCAACGCTGCGCATCTTCACTGAGCGCGTATCTATAGGGGCACTTTTGC 2142
DB 1040 GGTTTGAGAAAAAATCGCTGCGCTGATGAGATGCT-----GATAGACCATTTTGC 1094
QY 2143 CGTTATGCGCACCGCTCTACGCACTTTCACGATTCGACGCGCTGGCGCAACGCT 2202
DB 1095 GGTCAATTGACCGCGCTCTTACAGCGATGCGAGGCGCTTACGACACTTATGACGCGCTT 1154
```

QY	2203	GCTGGCGCGGACATGCGATCACTTTTATTCAGCAAGTATGACCGCACCTTGCTGAG	2262
Db	1155	AATGCGACGCGGACATCAGATTACTTTATCATCAGACGAGCAGTTAGCGGACTACAC	1214
QY	2263	CGACAGCGCAATTGCTTTGTGCGGTGGGAGGACCGAATCTGCGGATCGCTTC	2322
Db	1215	CGATAGCCGTAATCGGCTTTTTCGCGTGGTTTAAAGCTTCGATCCGCGGGCAGTCTGGC	1274
QY	2323	CAGCGAACTCAGGCGGCTGCGCCGACCGGGGCGGCTGTGCGCTGTTTCCGATTCACGA	2382
Db	1275	ACAACACCTTGCAATCGGGGCGGCACTCTCTCGGCCCCCAATGCTGAATGATCAATGA	1334
QY	2383	TCTGGCCGACGACCAACCATATATGCTATGCGCGAATGCTCGCGGTCTGCAACGCTGCA	2442
Db	1335	GATGGCCCGCAGACCCGATATGCTCTGTGATGAATGCTCCGNGGCTGACACCCCTGGC	1394
QY	2443	GCTGGAATGCGCTGAATGGCCGATCCAAATGGAAGCGGCTGTGTGTCTGTGTGGAGAGCTT	2502
Db	1395	GATTGATGCGGTGATGTGCAATGCAATGAACTGCGCTGCTACCTGGGCGGAGAAAGCCT	1454
QY	2503	ACAGGTGCGGTTCGTTGCTGCTGAGGCGCTGCGGCGCTGCGGCTCATGCGCAAGCGGCATTC	2562
Db	1455	GAACTGCTTATATGTTTGTGTTGCTGTGCTGTGCGCTGCTCACTTAAACGTGAAGGGAATTTCC	1514
QY	2563	GCTGTGTGTGATGCGCTTTTCGCTTTGTCTCAGAAATGAAAGCGCTGCACGCTATACAGC	2622
Db	1515	GCTGGCCGTAATGCGGTTTGACTATGCGCAGGACCGATGAGGACGGGAACGCTATGCGAC	1574
QY	2623	CAGCAGTGAATCTACGACCGCATATGCGGTGTATATGCGGCTGTCATGCTCTGTATGC	2682
Db	1575	CAGTGAATAAATCTATGACTGGCTGATGACAGTCAACATCGGGTGAATGCGCCGCACAGC	1634
QY	2683	GCGGCGCTTCGAGCCGCGCCGAAACGCGCATAGGCTTACATGATGTCTGTGCGCGCTGCGCA	2742
Db	1635	TCAATGCCATGGGAATTAGACACACGGGAAAACTGATCATCTGTTTTCACCGGTGGCGCA	1694
QY	2743	AATCAGTCAAGTGTGGCCCGCTTTTGAATTTTTCACGCGCAGCACTGCCAGGCTGTATCA	2802
Db	1695	GATCAACACGACTGATTCGGAACTGGAATTTTCCCGGCAGACATGCGCCGATTCATTCCA	1754
QY	2803	CAGGCTGGGTCCGCTGCGGAC----TTCAGTTGTAGCGGCGGCGCTTGCGGCACCTGG	2857
Db	1755	TTTCGGTTGGCCCGGCTGCTACCACTGAACCGGACCCCGCCCATCTCAGCTCGTACTT	1814
QY	2858	CCAGGCGCTGCGCCAG-CCGGTGTGTATGCGTCGCTGGGGAACGCTACAGGGGACATCGCTT	2916
Db	1815	CCCGGATGCGCACAGACCGCGCAATTTTGTCTTCACTCGGACCGCTGCAGGSAACGTGTA	1874
QY	2917	TGCGCTGTTCTGCACTCTGGCTCAGGCTGCGCGCAATCAGCAGCTGTGCTGTGTGTGGC	2976
Db	1875	CGGCTGTTTAAAAACAATCGTCCGGGCTTGCAGAGAAATGCATGCTCAGTTATGTGTGGC	1934
QY	2977	ACACTGTGGCGGGTTGACCGCCAGCCAGGCAATCAGTCAAGACTGTGCTGCTGCTGTG	3036
Db	1935	GCAATTTGCGTGTGTGTCACTTTTCAGGCGGAGAACTGCGCGCACGCACTGTTCA	1994
QY	3037	GGTACCGAATTTTGTGATTCAGCGGGGCGGCGCTGAGATGCGGACATGTTATCACTCA	3096
Db	1995	GGTGTGATTTTGTCCGACAGGCGGCTGCACTGGCGGAGCCGATCTTGTATCAATGCGCA	2054
QY	3097	CGCCGCTCTGAACAGTGTGCTGGAAGCACTGAGTGTGGACCGCCGATGCTGGCGTGGC	3156
Db	2055	CGGGGCGGATGAATACGTGTGCTGGAATTTAACCACTGACGCTGCTGTGAAGATGCC	2114
QY	3157	GATGCGCTTGATCAGCGCGGCGTGGCGGCACTGATTTGATGTGCGACGCGGCTGCGCGCG	3216
Db	2115	GCTGGCTTTTATCAGCCGGGTGTGTGCTGCAGAGTGTGCTGTGCGACGCTATCGGACGTG	2174
QY	3217	CGGCTCAAGTTTCAAGCGGGGTGCGGAGCTGGAAGCAACCTGCAACAGTTGTGAGTGA	3276
Db	2175	CGCTCAAGCTTTAACACAGCAGCACTCCATGTGACGTAGCTTCCAGACCTTAACTGTGCTGA	2234
QY	3277	CGATTCGCTATCGTGTGCGCATGTGAGCCATTCAAGGCGGACGCTGACGCGGCGGTTGCTG	3336

Db	2235	TGAAGCTACGTTACGAGAGATGAAACGCTGCGTAGCGCACCCTTGTCAGCGGCGGAC	2294
Qy	3337	TACGGCGCGGCTGATATTTGTCGAGCAGGCGCTGTGTACAGCAAGAAATCGTGTGCGGA	3396
Db	2295	CACCTGGCGGGGACATGTCGTAACAGGCGATCTGACCCGCAAGCCGGTGTCCACAG	2355
Qy	3397	GGCCACCTGATGCGCGCACCTTATGATGTCATTTCTGTGCGGTGCGGCTGTGAACGGG	3456
Db	2355	GAGACACTATGTCGGG----GTATGATCTGATTTCTGGTAGGTGGGGCTGGCCAAACGGG	2410
Qy	3457	CTGATTTGCGCTGCTTTACGCCACGTGACGCCGCACTTAAAGTTTGTCTACTGAGAGCT	3516
Db	2411	CTGATGCGCCCTGCGCCTGAGACACACAGCGGCGCTCCCTGCGCACTTGTGCTGATTTGACGCC	2470
Qy	3517	CAGGCGGACCGGCGCGCAATCAATCCGTGCTTCCATGCGGAAAGACGTGAGGAAAGCG	3576
Db	2471	GAACTGAAACCGGATGCCAATCAATCACTGGTGTTCATGCGGAAAGATCTTACGAAACG	2530
Qy	3577	CAGTTTGCCTGGCTGAGCGCGCTCTTTTCGCGCGCTGACCCGGGTATACAGTACGCTTC	3636
Db	2531	CAGCATGCTGTGATGCTCCGCTGTGTAGTACATCACTGGCCTGTGATGAGTCCGCTT	2590
Qy	3637	CCCAACCTGCGTGCACAGCTGATGATGTAATTTGCTGATTTGCTCGAGAAATTTTGC	3696
Db	2591	CCCCACGCAATCCGACGATGAACAGTGGCTATTTTGTGCGTACGCGAGACGCTTCGCG	2650
Qy	3697	CGGCATTTACAGCAGGTGCTCGGTGCGCGGTACGACCCGAGCGCGGTACAGGAGGCTC	3756
Db	2651	CAGGTCAATCCGACAGGTTTGGCCCGGATGTGCTGTAATTCACCGGTTTGCAGCATC	2710
Qy	3757	TCACCCACCGGGGTACAGCTGCGGATGCGCGGATGTTTACAGCGCAGCGGATTTGAC	3816
Db	2711	GCTTCAGCGCGGTCAAGCTGGAAGATGGCGGGGTGCTGAGAGTGAACGAGTATTTGAT	2770
Qy	3817	GGACGCGGGCTGCACCGCGACACCGCATCTGCACTCGCTATCAGGCAATTTGTGCTAG	3876
Db	2771	GGCGGTGCTACAGCCGATGAGCGCGCTGCGCATGGGCTTTCAGTGGTTTGTGCGACAG	2830
Qy	3877	GAGTGGCACTGGCGCGCGCGCATGCGCTGACAGCAATTTGATGAGACCGACGCTC	3936
Db	2831	GAGTGGCACTGCAGCGAGCCACATGGCTGACCGCGCGATCAATGAGTGAACCGTA	2890
Qy	3937	GATCAGCAGCAGGGTTATCGCTTTGTTTACACCTGCGCGCTCAGTGTGCAGCGCTTACTG	3996
Db	2891	GATGACGAGGAGGCTATCGCTTTGTCTACAGCTTGCCTTTTCTGACAGACACCTGCTG	2956
Qy	3997	ATTGAAGATCACCTACATCAACCATGCGCAGCTGTGATGCGGACAGGCGCGCTTAC	4056
Db	2951	ATTGAAGATCACCTACATGATTAACGCCACGCTGGAAGGCGATGCGGCGCTGACAGAT	3010
Qy	4057	ATTACGATTTATGCCACACAGCGCGGCTGGAATTTGGCCAGGCTGTGCGGAGAGACAC	4116
Db	3011	ATTCACGACTATTTGCGCGCAGACAGAGATGGGATTTAACCGGCTGTACCGGAAAGAA	3070
Qy	4117	GGCTGCGCGCATGACGCTCAGCGGGCGGATATCGATCAGTTCTGTGGCAACAGCAGCAGG	4176
Db	3071	GGCGGTTTACCATTTAGCTTACCGGGAAGCGTGGCGGCTTTCTGG---CAGAACCGGAT	3127
Qy	4177	CAACCGTACAGCGGAGCTGCGCGCGGAGCTGTTTACAGCCACACACCGTTACTGCTGCC	4236
Db	3128	TTGGCCGTGACAGCGTTTGGCGCGGGTTATTCATCCCACTTACCGGTTATTTCTGCA	3183
Qy	4237	GGCGCGGTGGCGCTGGCGGAGAAATTTGCCAGACGCTGCCCGGACGCTTCAACAGCTG	4296
Db	3188	CTGCGGCTGTCTGCGGTGACCGGCTGGCGGATGACAGATTTTACTTGAACCTGT	3247
Qy	4297	AGCCACTGCAATCGAATCTTTGACCGCTCAGCACTGGCGCGAGCGCTTTTTCGCTGTG	4356
Db	3248	CAGCGCACTTACAGCAATTTGCGCAAGTCAAGGCTTGGACAGAACAGCGCTTCTCCGATG	3307
Qy	4357	TTAATGCGATGCTGTTTCTTTCGCGGACGCGCTTGAACAGCGCTGGCGCTGATGCAAGCT	4416

Db	3308	CTGAACCGTAAAGCTGTTTCTTGAGGAGGATCCGAGCTGACCAAGCGCTGCAAGTTATGCAACGT	3367
QY	4417	TTTTTACCGGCTTGAAGCCCGCAATTGATTAAGCGCTTTTAAAGCCGGGCAACTGCGCTCAAGC	4476
Db	3368	TTTTTACCGGCTTCCCGAAGGTCATAGCCCGTTTATATGAGGAAAACTCACTTTGCT	3427
QY	4477	GATTAAGCAAGCATTTCTGTGCGGAAAAACCGCGAGTCCCTCTCGGCGAAGCGCTGCGCA	4536
Db	3428	GATGGGCTGGGCACTTTAAGGAGGCAAGCCCGGTCCCGTTCTGGGGGGGTTAACAGGC-	3486
QY	4537	TTGATGATGACCTCTCCGTTACCAAGGAGAAGATATGAAACGCACTTAATGTATGGCG	4596
Db	3487	--TATTATGACTCCCTCAACCGTCAACAGCGCATGCAATGAATGAATCAAGTATTTGGCG	3544
QY	4597	CAGGCTTCGGGAGCTGGCGCTGGCGGATTCGTCTGCAAGCGGCGGGGCTGCGTCAAGC	4656
Db	3545	CAGGCTTTGGTGGTCTGGCTTGGCAATTCGCTTCAGGCGTCAAGGCGTTCCACCCGAC	3604
QY	4657	TGCTGGAACAGCGCGATTAAGCTGGCGGCGCGGCTATGTGATCAGATCAGGGTTTAA	4716
Db	3605	TGCTGGAGCAGCGTGAACAAGCTGGCGGCGGGCTTATGTATCAGGATCAGGGCTTCA	3664
QY	4717	CCTTGAATGCGCGGTCCGACGGTGAATTAACCAATCCCAAGGCTATGAGGCGGTGTTACGC	4776
Db	3665	CGTTTGATGCGGCGCCCAAGGTAATCAACGATCCGAGCGCATTAAGAAAGCTGTCAACC	3724
QY	4777	TGCGACGCAAGCAACTCAGTATTAATGTCGACTGATGCGCGTGAAGCCATTTATGGC	4836
Db	3725	TGCGCGGTAATAAAGCTCTGACATATGTGAGCTGATGCGGTGAAGCCGTTTATGGCC	3784
QY	4837	TGTGCTGGGAAGAAGCGCAGCGCACTGGACTTACACACATATCAGCGCAGCTGAGCAAGC	4896
Db	3785	TCTCTGGGAATCGCGCAAGGTGTTCAGTTATGACAAAGATCAACCCGCGCTGGAAGGCG	3844
QY	4897	AGATTGCACTTTTAATCCCGAGAGTGTGCGCGGTTACCGCAATTTCTGAGCTATTACAC	4956
Db	3845	AGATTGCGCGCTTTATCCGCTGACATGTTGAAGATATACGTTCGCTTCTGGCCATATCCC	3904
QY	4957	AGGATGTGTTTCGTGAGGGCTATCTGAAACTGGGCAACCGTACCTTTTCTGCAATTTCCGCG	5016
Db	3905	GAGCGGTTGTTGCTGAAGGCTATCTGAAGCTTGGGCAACGTCGCTTCTGTCAATTCGCGC	3964
QY	5017	ACATGCTGCGGTGCGCGGCAAGCGTGGGTGGCGTGCAGGCGTGGGCGAGTGTACAGCA	5076
Db	3965	ACATGCTGCGGCGCGCGCTCAAGCTGGCAAACTTCAGGCGGTGGGCGAGCTTTTACAGCA	4024
QY	5077	TGCTGCGCAAAATTTATTCATGACGATCACTTGCAGCGACTTTTCTTCACTCGTTGC	5136
Db	4025	AAGTGGCGAGTCACTTAAGTATGACATCTGCGTCAAGCCCTTCTTTCACATCACTGC	4084
QY	5137	TGTGTCGGCGGTATCTTTTGGCAAGCTTGTGATTAATCTTAAATTCAGCACTGAGAC	5196
Db	4085	TGTGTCGGCGGAAATCTCGTTTGCACTTCTCAATCTTAATCCTGATTAATCGCTGGAAAC	4144
QY	5197	GCGAATGGGGGCTGTGTTTCGCGCGGCGGTTACCGGTGCGCTGGTTGATGGATGGCGC	5256
Db	4145	GTCGAATGGGGGCTCTGTGTTCCGCGCGGTGGCAACGGGCGCGCTGTGTGACGGGAAATGGGA	4204
QY	5257	GAGCTGTTTCGCGATTTGGGCGGTGAATCTGTGCTCAACGCCGAAATGACCACTGCGAGA	5316
Db	4205	AAGCTGTTTGAAGATCTGGGCGGCGGAATGGAGCTCAATGACCAAGCGTTCGCCGCTGGAGA	4264
QY	5317	CCGAGGGTAACCGCATCAAGCGGTGTCCAGCTGGAAGAGATGGGGCGCGTTTGGCGCGCGCG	5376
Db	4265	CCGAGGAAACAGATTTACCGGGGTGCACCTGAAGATGGCGGGGTTCTTCCGACCCGCG	4324
QY	5377	CGGTTGCGTCAAAATGCTGACGTGTGCAATCCTACGATCGCCTGTTAAGCAGCAATCCTG	5436
Db	4325	CGGTTGCGTCAAAAGCAATGTGTTCACCTTACCGGAATCTGTGAGCAGCATCCCG	4384
QY	5437	CGGCGCGTAATACGCGCGCAACGCTGAAGCGCAAGCGGATGAGCAATCTCGCTGTTTGAAC	5496
Db	4385	CTTGCACAGCGCAGGAGCAATCACTGCAAAACAAACGATGAGCAACTCACTGTTTGAAC	4444

QY	5497	TCATTTTGGCTTAATCAAGCCACCCGAGCTGGGAGCACACAGGATGCTTTGGTC	5556
Db	4445	TCATTTTGGCTTAATCAATCAATCAATTAAGCTGGGAGCACACAGGCTGCTTTGGTC	4504
QY	5557	CGCGCTATCGTGAATTGATCGATGAGATCTTCAATAGCAGCCAGCTGGCGAAGATTCT	5616
Db	4505	CGCGCTATCGGAGTTGATTGATGAGATCTTTAAACAAGATGGCTGGCAGAGACCTTCT	4564
QY	5617	CGCTGATATCGCATATGCCCTCTGCTCCAGAGCATCCGCTCGCTGGGACCGGCGGGGCTGGGGCA	5676
Db	4565	CGCTATATCGCATATGCCCTCTGCTGGTGCAGATCCCTCACTGGGCGCGAGAGGCTGGGGCA	4624
QY	5677	GTTTTTCAGTGGCTGGCCCGCGGTGCGCGATCTCGGTACCCGCGCAATTTGACTGGCAACAG	5736
Db	4625	GCCTACTAGTGGTGGGGCCAGTTTCGGACCTCGGACCCGCGAATTCGACTGGGCGGTTG	4684
QY	5737	AAGGCGCGCTTGGCGCGATTCGCATTTTGTCTTATCTGAGAGAGCATATATCCCGGTC	5796
Db	4685	AAGGTCCGCGCCTGGCGGATTCGATTTTATACATCTGGAACAGCATATATTCGCGGC	4744
QY	5797	TGCGACAGCAGTTAGTACACACCGTATGTTTACCGCTTTGATTTTTCGCAACAGCTGC	5856
Db	4745	TGCGTACCGAGTTGGTACCGATTCGATCTTCAAGCGCTTTGATTTCCGGAGTAGAGTGA	4804
QY	5857	ACGCGCATCAAGGCTCAGCGTTTTCGCTCGAACCCATTTTACGCAAAAGCGCTGGTTC	5916
Db	4805	ATGCGTATCAAGGTTTCGGCTTTCGGGTGAGCGCATCTCGACGCAAAAGCGCTGGTTC	4864
QY	5917	GGCGCGCATTAACCGGATATGCCGACATTAATCTTAATCTGGTGGGGGCTGGGACGATC	5976
Db	4865	GGCTCTCAACCCGGATPAAAAATTAATATCTCTATCTGGTGGGTGAGGACGGCATC	4924
QY	5977	CCGATGCGCGGTGTGCAGAGCGTGAATCGGCTCCGCGAAGCGAACCGCCAGCTGATGGTG	6036
Db	4925	CTGGCGGGGATTCAGAGGGTAAATTTGGCTGGCGAAGGCTACCGCAGATTAATGGCTGG	4984
QY	6037	AGGATCTTGACCGGATGAACCAACCGCGCTGATTTGACAGAGTCAGCAATCCATGGCGCA	6096
Db	4985	AGGATCTG---GCTTGAATATGTCCGTACCTGCTCGATCAAGCCGTAGACACCATGAGGT	5041
QY	6097	GGGCTCCAAAAGTTTGGCGCAGCGCTACCCCGGCTAATTTGATCTCTTCAACGCGCGCAGTAC	6156
Db	5042	GGGATTCGAAAAGCTTTGGCCACCGCGTCAAAACGTGTTGGTGCMAAACCCGACGAGCGT	5101
QY	6157	GCTGATGCTGTACGCTGTGTGTCTCACTGTGACGATGTATAGATGTCAAGCGCTGGG	6216
Db	5102	GCTATGCTCTACCGCTGTGTGCGCTCACTGTGATGATGTATTTGACGATCAGGTACTGGG	5161
QY	6217	CGAAGCGCGCACGAGCACCGCGGTGGGAGATGCAACAGCGCGGATGGCGCACCTGCACAT	6276
Db	5162	ATTGACGACACATACGCGCATGCTGTGCAATCCGCCGAACAGGCGCTGGCGAGCTGGAGAT	5221
QY	6277	CGAAACCCGCGCGCCTCTACAGCGGTGGCCAATAGATGAACCAAGCGTTTGTGTCTTTCA	6336
Db	5222	GAAAACGGTTCAGGCGCTATGCGGATCCGAGATGATGAGCGCGCTTTGCAAGCTTTTCA	5281
QY	6337	GGAAATGGCGCTGACGATCAAGCTTCCCCACAGCTGCGCTTTTGTATATCTGGAAGGTT	6396
Db	5282	GGAGGTGGCATATGCAACGATATTTCGCTGCTTACGCTTTGTATATCTGGGGGCTT	5341
QY	6397	TGGCATGGATCGGGTGAAGAACTTATAGCGTGTTCGGGGAACACGCTGCGTTATCTCTA	6456
Db	5342	TGCGATGACGTGATAGACACGCTTATCAGAAGCTGAGATGATAGCGTTCGGTTACTGTTA	5401
QY	6457	TCAAGTGGCGCGCTGTGGGTTAATGATGGCGCGGCTGATGGGCGTACGTGATGAGCG	6516
Db	5402	CCAAGTGGCGGGCGTGGTGGCTGATGATGGCGCAATATATGGGCGTTACGCAACAGC	5461
QY	6517	CGTACTGATATACGCTCTGTATTTGGGTCTGGGCGCTTCAAGTTAACCAATATCCGACGGGA	6576
Db	5462	CACGCTGGATGGCGCTCGCATATTCGGCTGGGCTTTCAAGTGTGCAATATATGGCGCGGA	5521

```

QY 6577 TATCGTTGAGACAGCGGAGATGCGGTTGATCTGCCAACAAGCTGGGTGATGAGGC 6636
DB 5522 TATCGTTGAGATGCTGACGCGGAGGCTGCTATCTGCCCTACCTGCTGCTGTAAGA 5581
QY 6637 CGAAGTACGCGCGCGCCAGCTTGCATGCGCAATCGCGAGCGCTGAGCCCGCTGCG 6696
DB 5582 GGGGCTTACCCAGAGAAATCTGCCCATTCGCAAAATCCGACAGCATTAAGCCGCTCG 5641
QY 6697 AGCGGCTGTGTGCGGAGCGCGAGCCGTAATATCATGTCAGCGCGAGCGGCTGCGCAG 6756
DB 5642 CGCGCGGCTGTGTGAGAAACGACAGAGCCCTATATCATGCGCGCTGCGCTGCGCGG 5701
QY 6757 ATTGCGGCTCCGTTGCGGCTGCGGCGATTCGCAACCGCGCGCGGCTTACCGGAAATTTG 6816
DB 5702 TTATCGGCTGCGCTTACGCGTGGGCGATTCGCAACCGCGAGCGGTCTACCGTAATTCG 5761
QY 6817 CATAAAGTGCAGCATGCGGCTGCGCGGCGATGGAATACGCGCGACGCAAGTAAGG 6876
DB 5762 TATGAAAGTGTGTGACGCGGCTTCAAGCGGCTGGAATCAACGCGCATGCCACGACACC 5821
QY 6877 CGAAAAGCTGGCGCTGCTGTGTAAGGTGCGCGGCTTACTTTCGCGCTTGTCTCA 6936
DB 5822 AAGAAATGCGCACTGCTGTGCGGCGATTCGCGGTCAAGCGGCTTACTTCCGCGGTGCGG 5881
QY 6937 TCCCGAGGCGGCTGCTGCGGCTGTGCGAGCGTCCGCGTTCAGACGCGCATGCGCG 6996
DB 5882 TCACGCTTCACGCTGCTGCTGATCTCTGGAAGCGCGCGCTTAAAC-----GCCATACGT 5935
QY 6997 TGGCGAGGCTGCGCTGCGAGCTGTGCAACGCGTGGGCGCTTAAGAAAGCCAAAGAAAG 7056
DB 5936 TCTCGAGGCTGCGCTGCGATTTTGAACGCGGCGCGCGCTTAAGAAAGCCAAAGCAAGCA 5995
QY 7057 CAGCCTTCCGCTGCGCGACCGCGATGATGATGCGGCTGCGCATATCAACCGCTTCA 7116
DB 5996 CAGCCTTCTTGGCTGCTGCTGATCCGCGATGATGATGATGATGATGATGATGATGATG 6055
QY 7117 TAGCCTTGGCTGATATATGCGAAACGCGCGCTTATGATGATGATGATGATGATGATG 7176
DB 6056 TAACCCGCGCGCGCAATAGCGGAATGCGGCGCTGCGACGCGCGCTGATGATGATG 6115
QY 7177 ATAAATAGAGAGCGCATAGTGTGATGCGCGCGCAATTCATGCGAGCGCGCGAGAG 7236
DB 6116 ATAAATAGAGAGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 6175
QY 7237 CCGTTGACACCCAGCCAAATCAGCGCATGCAACAGCGCAACCAACGCGCATAGG 7296
DB 6176 CCGGCTGCGCGCGCAATATGAGCAAAATGCAACAGCGCGCAACCAACGCGCATAGG 6235
QY 7297 TCGTTAGCTCAAAATTTGCTCTCATGTGTTCAATGCGCAAAATGCGCGCATGCC 7356
DB 6236 TCATTAACCTCAAAACGAGCTTATGCGGTTCAATGATGATGATGATGATGATGATG 6295
QY 7357 CAGCCATGATATATGATTTATGCGACAGCGCGCGCGCATTCATGATGATGATGATG 7416
DB 6296 CAGCCATGATATATGATTTGCTGCGCGAGCGCGCGCTTATGCTGCTGCTGCTGCTG 6355
QY 7417 GCCAGCAGATTAAGCCCATTCATTAACAGAGCATATGATGCTC 7460
DB 6356 ACCAGAAACATCAGGCGATTCGCAATCCACACATATGTTCTCC 6399

```

```

RESULT 2
LOCUS AY166713 6586 bp DNA linear BCT 18-DEC-2002
DEFINITION Pantoea stewartii zeaxanthin-diglucooside biosynthetic operon,
complete sequence.
ACCESSION AY166713
VERSION AY166713.1 GI:27228289
KEYWORDS
SOURCE Pantoea stewartii
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Pantoea.

```

```

REFERENCE 1 (bases 1 to 6586)
AUTHORS desouza,M.L., Kollmann,S.R. and Schroeder,W.A.
TITLE Carotenoid Biosynthesis (WO 02/07395 A2)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 6586)
AUTHORS desouza,M.L., Kollmann,S.R. and Schroeder,W.A.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-2002) Biotechnology Development Center, Cargill
Inc., P.O. Box 5702, Minneapolis, MN 55440, USA
FEATURES
source
1..6586
/organism="Pantoea stewartii"
/mol_type="genomic DNA"
/strain="ATCC 8200"
/db_xref="ATCC:8200"
/db_xref="taxon:66269"
169..6399
/gene="zeaxanthin-diglucooside biosynthetic operon"
169..1080
/gene="crtE"
169..1080
/gene="crtE"
/codon_start=1
/product="geranylgeranyl pyrophosphate synthase"
/protein_id="AAN85596.1"
/db_xref="GI:27228290"
/translation="MVCCKKRVHLTGISAQLADLIDSLDQLLPVQGERDVCAGAM
REGTAPGKRIIPMLILATRDGCAISHGILIDACAVENYHVASLIDDPKMDMA
QMRGRPTHTGHEVAILAVALLSKAFGVIAEGLTPIAKTRAVEISLTATGMO
GLVQGFQDLSGDKPRASADATILTNOPFTSLFPGSTOMASIAANSCAEARNLRF
SLDLCQARQLDLDIDGKDTGKDTGKDTGKDTGKDTGKDTGKDTGKDTGKDTGKDTGKDTG
SAACNGHSTQLFIQWPFKLAAS"
1090..2385
/gene="crtX"
1090..2385
/gene="crtX"
/codon_start=1
/product="zeaxanthin glucosyl transferase"
/protein_id="AAN85597.1"
/db_xref="GI:27228291"
/translation="MSHFAVITAPPPSHVRLATONLAQELVARGHRTVPOOHCKALV
TSDIGFQVIGLTHPPGSLSHLILAHPLPSPMLRLNEVARTSDMLCRELPAAH
ALQIGVIVDQMEPAVAEASGLPFSVACALPLNRPGLPLVMPREYGISDAR
ERYTTSKIVDMRMRRHDEVIHNAORMGLAPREKLNCFSPLAQISQDLPIDLPK
ALPDCEHAGVPLROPQGTGPGSTSYSPSPDKPRIFASLTGLQHRGLFRTIAKCE
VDAOLIAHCGGLSATQGLARGGDIQVDPADQSAALSOQLTTHGAMNTVDLAI
KSTPLIAPLAPDPGVASRIIVYIGIGKRSRFTTSHALAQISLINTIYPOMT
KIQAURLAGTFFPADIVDQMRITCQPVLSGDIATAT"
2369..3517
/gene="crtY"
2369..3517
/gene="crtY"
/codon_start=1
/product="lycopenase"
/protein_id="AAN85598.1"
/db_xref="GI:27228292"
/translation="MQPHYDLILVGLANGLIALRLQOQHPDMRLILEAGPAGN
HTMSPHEDTLNQRWTAPIVHMHPDQVRFQRRRVNNGYCVTSRHPAGILRO
QFGQHLMTHTASVAHABSVQADGRIHTASTVILDRGTTPSALRFVGOAFLGQMO
LSAPGLSSPILIMDAIVDQNGYRFVYTLPLSATALLIDHTYIDVANIQAERAKNI
RDYARQGWPLQTLREBQALPITLTGDRQFQWQDQPCGSLRAGLHPFTVGLP
LAVLADLRSLADVFTSSVHOTIAHFAQOQCGFFMLNRMFLTPALPASRWBM
GRFYGLPEDLIARFYAKLTVDRILISGKRPVPFALQAIMTTHR"
3529..5007
/gene="crtI"
3529..5007
/gene="crtI"
/codon_start=1
/product="phytoene desaturase"

```

```
/protein_id="AA085599.1"
/db_xref="GI:27228293"
/translation="MKFTYVAGFGGLALAIRLQAGIPVLLIEQRDKRGRAVYVQ
EQGTFDAGPVTIDPSAIBELFALAKQLDVLLPTVPFRLCWESEKVFYND
QAQLEAOIQOFNPRDVAQYRAFLDSRAVNEGTLKGTVPFLSKMLAAQPLATL
QAMRSYKVAQYIEDEHLROAFSFLVAGNFPATSSYTLTLHALEREQWAFPG
GTGLVNGMKLFDGLGGEVYNARVSHMETVGDKIYVOLDERRETCVAVSNAY
VHTYRDLISQHPAAKQAKYLSKRMNSLFLVYEGANHHDQALHHTVCPGRYREL
IHEIFNHDGLAEDSLYLHAPCTVDPSLAPEGCSYVLAAPVHLGTANLDMVAGEER
LNRDIFDLEOHMPLGRSOLVTHRMFTPDPRDELNAWQSGASVSPILITQSAWFR
LNRDKNIDNLYLVAGTHPGAIGPVIGASAKATGLMLBDLI"
5043..5933
/gene="crlz"
/contig="5933"
/contig_start=1
/contig_end=1
/codon_start=1
/codon_end=1
/product="phycoene synthase"
/protein_id="AA085600.1"
/db_xref="GI:27228294"
/translation="MAVGSKSFAATSLFPDKTRRSYLMILYANCRHDDYIDDTLGF
HADQPSQMPBQRLQLEKTRQAYASQWHEPAAFOEVAAMHDTAPAYADHLEG
FAMDVRETRLTLDLTRYCYHAGVANGYLMMAQIMGRDADNLDRACDLGLAQLNI
ARIVDAQVRCYLPESMLEBGLTANYAAPENRQALSRIGRLVREAEPPYSSM
AGLAQLPLRSAMALATAKOYRKIGVKEQAGKQAMDRQSTAEKLTLLTASGA
VTSRMKTYPPRPALMORPI"
complement(5872..6399)
/gene="crlz"
complement(5872..6399)
/gene="crlz"
/contig="5933"
/contig_start=1
/contig_end=1
/codon_start=1
/codon_end=1
/product="beta-carotene hydroxylase"
/protein_id="AA085601.1"
/db_xref="GI:27228295"
/translation="MLWIMNALIVFVTVGMEVVAALAHKYLINHGWMGMLSHHEPR
KGAENVDLAAVFAIVSIALIYFGSTGIGLPIWQIMGAMVAGLYLWYMDVGHVQHW
PFRYIPRKYLKRLYMAHRMHNAVRGEGCVSRFLYAPPLSLQALRLRRHARSGA
ARBDQBDVDRSSSGK"
ORIGIN
Query Match 31.0%; Score 2688.4; DB 1; Length 6586;
Best Local Similarity 65.3%; Pred. No. 0;
Matches 4077; Conservative 0; Mismatches 2136; Indels 27; Gaps 8;
QY 1257 CAGTTGCATACGCGATTGACAGGCCCATCTTGAACATTATTGCTCCGCGGACAGAGCC 1316
DB 220 CAGTTGCTGCTGATATCGATAGCCGCTTGATCAGTTACTGCGCGGTTCAAGGGTGAAGCG 279
QY 1317 GATCGGCTTGGGCGCGCATGCGTGGCGGACCGCTGGACCGGCGCAACGTAATTCGTCG 1376
DB 280 GATTGTGTGGGTGCGGATGCGTGAAGGACCGCTGGACCGGCGCAACGTAATTCGTCG 339
QY 1377 CTCTTGTGCTGTCGACGACGCGATATGCGCTGTGACGTGGCGACGAGGCGATCCTT 1436
DB 340 ATGCTGCTGTATTATTAACGCGCGCGATCTTGCTGTGCGATCAAGTCAAGGGGATTACTG 399
QY 1437 GATCTTCCCTGTGGGCGTGAATGTGACGCTGCTCACTGATCTTCGACGACATTTCCA 1496
DB 400 GATTTACCTGCGCGGTTGAATGTGTCATGCTCCCTCGCTGATTCGATGATATGCC 459
QY 1497 TCATGTGATTAACGCGCGGATGCGACGTTGGGCGCGCGCGCATTCACCTGTGAATATGGGAA 1556
DB 460 TGCATGACGATGCGACGATGCGTGGGCGGCTGCCACCATTCACACGCGAGTACGGTGA 519
QY 1557 AACGTGCGATCTTGGACGCGTGCCTACTACGCGCGCTTGAAGTGAATGCCCTC 1616
DB 520 CATGTGCGATTTCTGGGCGGCGTGGCTTTACTACGAAAGGTTTGGGGTGAATGGCGAG 579
QY 1617 GCGCGGGGTCTGCCAGCAAGCAAAAGCGAAGCCATTGCGCAAGCTTCTCTGCGCG 1676
DB 580 GCTGAAGGTCTGACGCGATGACCAAAACTGCGCGGCTGTGCGAGCTGTCACTGCGATT 639
```

```
QY 1677 GCGCTGAGAGGACTGTGTTTCAGGGTCACTTCCAGATCTGATGACGCGGACACAGCGC 1736
DB 640 GGCATGACGAGGCTCTGTTTCAGGGCGAGCTTTAAAGACCTCTCGAAGCGATTAACCCGC 699
QY 1737 AGTCGGAACCATTCACCTCGAACCAATGAACCTGAAACACAGCTCTCTTTCGCGCAGC 1796
DB 700 AGCGCGCATGCCATACCTGCTCAACCAATCACTTTAAACGACGACGCTTTTGGCGCTCA 759
QY 1797 CTGCAATGCGCGGATGTCGCGCGCATGCTGATGAGTACGTCAAGCTTAAGCTAT 1856
DB 760 ACCCAATAGCGCTTCATGCGCGCAACGCGCTCTGCGAAGCGGTAAGAACTGCATCGT 819
QY 1857 TTTGCGAGAGATTAGGTCAGGCTTTCAGCTTCTGAGACGACTGGCGGATGGCTCTAG 1916
DB 820 TTCTCGCTGATCTCGGCGAGGCTTTCAGTTGCTTGAAGATCTTACCGATGCGATGACC 879
QY 1917 CACACCGGCAAGGACTGTCTACAGATTCAGGCAAAATCCACGCTGTGTGACATGTGGC 1976
DB 880 GATACCGGCAAAACATCATCATGAGATGACGTAATCAACGCTGGTCAATTAATAGC 939
QY 1977 CCGAAGGGGCTGAGCGGTCTGTGCGGACATCTTAAGACGCGCGATGACACCTTGGC 2036
DB 940 TCAAGCGCGGTCGAAGAACGCTGCGACAGCATTTGCGCTGCGCAGTGAACACTTTC 999
QY 2037 TGGCGCGCATGCGG---TGTGCGCACCCGCAATATATGACGCGCTTTAATCA 2093
DB 1000 GCGGCAATGCCAAACGCGCATTCACCAACCACTTTATTCAGGCTGTGTTGACAA 1059
QY 2094 CAGCTGCGATGTCTTCACTGAAGCCGCTCATACCTATGGGCGATTTTGCCTATTTCGC 2153
DB 1060 AAATCGCTGCGCTGATTAAGATGT-----GATGAGCATTTTGGCGTATGCGAC 1114
QY 2154 CAGCGCTGACAGCCTTTCACGATTTGAGGCGCTGGCGCAAGCGTGTGCGCGC 2213
DB 1115 CGCGCTTTTACGCAATGTTGCGCTCTGCAAAACCTTGTCAAGATTAAGTGGCGCG 1174
QY 2214 GATCGCATCACTTATTCAGCAAGATGATCAGACGACTTGTGAGGAGCAAGAGCG 2273
DB 1175 GTATGTGTTAGCTTTTTCAGCAATGATCTGCAAGAGCGCTGTAAACGAGCGAT 1234
QY 2274 TTGCTTTGCGCGTGCAGAGCGACGACGATCTCGAGATGCTCTCAGGAACTCA 2333
DB 1235 TCGGATTCAGACGATCGAGCTGCAAAACGATCTCCGCGTCTTATGCACTGCTG 1294
QY 2334 GCGGCGTGGCGGACCGGCGCGGCTGTGCTGTTTGTGCTGATTCAGTCTGCGACGA 2393
DB 1295 ACCTGCGCGGACCCACTCGAACCTCTCATGTTACATGATCAATGAATGGACGA 1354
QY 2394 CCAACGATATGCTATGCGCGGAACTGCGCGGCTGTGCAAGCGCTGCAAGGTGATGGCG 2453
DB 1355 CAGCGATATGCTTTGCGGGAATCTGCGCGCTTTTCAATGCTTGCATGATGAGGGCG 1414
QY 2454 TGATTGCGATCAATGGAAGCGGTGATGCTGTGTGCAAGAGGTTACAGCTGCGT 2513
DB 1415 TGATGCTGATCAATGAGCGGAGCGGAGGTGAGTATGCAAGAAAGCTGCGTCCGT 1474
QY 2514 TCGTGTGCGTGTGCGGCGGCTGCGGCTGATATGCGAAGCGCATTCGCTGTGTGTGA 2573
DB 1475 TTGTTTTCGTTGCGGCTGCGGCTGCGCTCAACGCGCAACCGGGTTTGCCTTGGCGTGA 1534
QY 2574 TGCCTTTTGTGCTTGTGCGAGATGAGAAAGCGCTGAGCGCTATACGCGCAGAGTGA 2633
DB 1535 TGCCTTTTCAATGAGGACCAAGGATGCGGCTTGGGAAGCTTATACCAACAGGAAATA 1594
QY 2634 TCTACAGCGCATCATGCTGTGCTGATGCGCTGTCAATGCGCGGCGCTTGC 2693
DB 1595 TTATGATGCGGCTGATGCGACGTCAGATCGTGTATGCGGATATGATGAGAAATGG 1654
QY 2694 GCTTGCAGGAGCGCATGCTTATCATGATGTGTGTGCGCGCTGGCGCAATCATGTCAG 2753
DB 1655 GTTTACCGCGCGGTGAATAAATGCAATCATGTTTCTTCACTGGGCAAAATTCAGCAAT 1714
QY 2754 TGGTGC CGGCTTTTGAATTTTCAAGCGCAACATGCGACGCTGTATCAAGGCTGGCT 2813
```



```
Db 1715 TGAATCCCGAAGTGGATTTTCCCGCAAAAGCTGCGCAAGCTGTTTCAATGCGGTTGAC 1774
QY 2814 CGCTGGGAGATCCAGATTGCTAGGCGCGCTCGCCGCAACCTG---GCCAGCGTGGCC 2870
Db 1775 CGTTACGGGCAACCCAGGGGAGCGCGGGGATCATCACTTCTATTATTTCCGTCGCCGAGCA 1834
QY 2871 AGCGGATGGTATATGCTCGCTGGGCAAGCTACAGAGGAGCATGCTTTGCGCTGTTCTGC 2930
Db 1835 AACCCGATATTTTGGCTCGCTGGGCACTCTGAGAGGAGCATGCTTATGCGCTTTACAGA 1894
QY 2931 ATCTGGCTAGGCGCGCGCAATCAGAGCTGTGCTGCTGGTGGGCACTGTGGCGGCT 2990
Db 1895 CCATCGCCAAAGCTCTGGAAGAGTGGATGCGCATTTACTGTTGGACACTGTGGCGCC 1954
QY 2991 TGAACCGCAGGAGGAGCATCAGCTCAGCTGCGCGGCTGCTGGGAGTGGATTTTG 3050
Db 1955 TCTCAGCAGGAGGAGGAGTGAATGCGCCCGGGGCGGAGCATTTAGGTTGTGGAATTTTG 2014
QY 3051 TGGATCAGCGGCGCGGCTGAGAGATGCGCACTGTTTATCACTCAGCGCGTCTGAACA 3110
Db 2015 CCGATCAGTCCGAGACATTTCAAGGCAAGTGAACATCACAATGATGAGATGATA 2074
QY 3111 GTGCGTGGAGGAGTGGAGTGGAGCAAGCGGAGTGGGCTGGCGGATGCGCTTGATC 3170
Db 2075 CGGATCTGGAGCTTATTTGCTTCCCGCACCGCTACTGCGCTGCGCTGGCATTTGATC 2134
QY 3171 AGCCCGCGTGGCGGAGCATATTTAGTGGAGCGCGCTGCGCGCGCGCTCACTGTTCA 3230
Db 2135 AACCTGGCGGATCAGCAATTTGTTTATCATGGCAATCGGCAAGCGGCTCGCTGGTTA 2194
QY 3231 GCGGCGTGGCGAGCTGGAGACCACTGCAAGTTCGAGTGGAGATGCTATGCTC 3290
Db 2195 CTACCAAGCATGCGCTGGCGCGGAGATTCATGCTGCTGACTAACCGATTAACCCG 2254
QY 3291 TGGCAGTTCAGGCAATTCAGGCGGAGCTGAGCGGCGGCTGAGCGGCGGCTG 3350
Db 2255 AGCGTATGACAAATTCAGGCGGCGCATTTGCTGAGGAGGAGCACCGCGCGCG 2314
QY 3351 ATATTTGAGAGGAGGCTGTGTCAAGAGCAATTCGTGCTGGCGAGGCGCACTGATGC 3410
Db 2315 ATATTTGTAAGAGGAGGATGCGGACCTGTCAAGCAAGTACT-CAGTGGGAGAGTTATGCA 2373
QY 3411 CGCACTTATGATGCTATTTGCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 3470
Db 2374 ACCGCACTATGATCTATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2433
QY 3471 TTTAAGCGAGTGGAGCGCGGCACTTAAGTTTGTCTACTGAGAGTCAAGGCGGAGCGG 3530
Db 2434 GCTTCAGCAACAGCATTCGGATATGCGGATCTTGTATTGAGGCGGCTCTGAAGCGG 2493
QY 3531 CGGCAATCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3590
Db 2494 AGGGAACATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2553
QY 3591 CGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3650
Db 2554 AGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2613
QY 3651 CGAGCTGAGTGGATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3710
Db 2614 CATGTGAACAGTGGCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2673
QY 3711 GGTGCTCGGTGCGCGCTAGCAACCGGAGCGCGGCTGAGGAGTCTCAACCGGAGT 3770
Db 2674 AAGTGTGAGCAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2733
QY 3771 CAGATGGCGGATGGCGGAGTGTAAAGCGGAGCGGCTGATTTAGCGAGCGCGGCTGCA 3830
Db 2734 CCGATTAGCGGATGGCGGAGTGTAAAGCGGAGCGGCTGATTTAGCGAGCGCGGCTTAC 2793
QY 3831 GCGGACAGCGCATCTGAGCTGCGCTATGAGGATTTGCTGCTGAGGAGTGGCACTGCG 3890
Db 2794 GCCTGATCTGCACTACCGGTAGGATTCAGGAGCTTTATCGTCAAGAGTGGCACTGAG 2853
QY 3891 CGGCGGCAATGCTGAGCAGGCAATATTATGAGCGCCAGCGTGTATGAGCAGAGG 3950
Db 2854 CGGCGGCAATGCTGATTCGTCACCGATTATCATGATGAGCAGCGTGTATGAGCAAAATG 2913
QY 3951 TTAATGCTTTGTTAATGCTGCGGCTGAGTGGAGCGGCTTATGATTAAGATACCA 4010
Db 2914 CTACGCTTTGTTAATGCTGCGGCTGAGTGGAGCGGCTTATGATTAAGATACCA 2973
QY 4011 CTACATCAACATGAGCAGCGTGTATGCTGAGCGGCGGCTTATGATTAAGATACCA 4070
Db 2974 CTACATGAGCAAGCTTATGCTGAGCGGCGGCTTATGATTAAGATACCA 3033
QY 4071 CACAGCGGCTGAGATTTGCGCAAGCTGCTGCGGAGAGCAGCGCTGCTGCGAT 4130
Db 3034 TGGCGCAGAGGCTGAGCGGCTTATGAGCGGCTGCTGCGGAGAGCAGCGCTGCGAT 3093
QY 4131 CAGGCTCAGCGGAGTATGATGATGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4190
Db 3094 TACGTTAAGGAGATATGCTGAGTGTGAGCAAGCAAC---CGCAAGCTGTAGCGG 3150
QY 4191 GCTGCGGCGGAGCTGTTTCAAGCGCACCGGTTACTGCTGCGCGCGGCTGAGCGCT 4250
Db 3151 ATTAAGCGCGGCTGTTTCAATGCGCAACCGGCTACTGCTGCGCGGCTGAGCGCT 3210
QY 4251 GCGGAGAAATTTGCGAGACGCTGCGCGCGGCTCAACAGCTGAGCCATGATGCA 4310
Db 3211 GCGGATGCTGCTGAGCGCGGCTGAGTGTATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 3270
QY 4311 ATCTTTGCGGCTGAGCAGCTGAGCGGAGAGCGCTTTTCCGCTGTTAAATGCGATGCT 4370
Db 3271 TCACTTTGCGGAGAAAGTTGGAGAGCAAGGAGTGTTCGATGCTGATGCTGATGCT 3330
QY 4371 GTTCTTGGCGGAGCTGTAACAGCGCTGAGCGGAGTATGAGCAGTATTAACCGCTTGA 4430
Db 3331 GTTTTAGCGGAGCGGCTGAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3390
QY 4431 CGCGGAGTGTATGAGCGCTTTTACGCGGAGCACTGCGCTCAGCGATTAAGCAGCAT 4490
Db 3391 CGAGATTTGATTTGCGCGCTTTATGCGGAGAAATCAACGTAAGCAGTGGCTAGCGAT 3450
QY 4491 TCTGTGCGGAGAAACCGCGGCTGCTGCGGAGAGGCTGCGGAGTATGATGATGATC 4550
Db 3451 TCTGAGCGGAGAGCGCGGCTTCCGCTTTCGCGGAGTGGAGCAATTAAGCACTCA 3510
QY 4551 TCGGTTACAGGAGAAATTAATGAACGCACTATGATGATGAGCGGAGCTGCGTGGC 4610
Db 3511 TCGTTAAGAGCACTA---CATGAACCAACTAGGATATGATGAGGCTTGTGCTG 3567
QY 4611 CTGCGCTGCGGATTTGCTGCAAGCGGCGGCTGCGGCTGCAAGCTGCTGAGAACAGCG 4670
Db 3568 CTGCGCTGCGGATTTGCTGCAAGCGGCGGCTGCGGCTGCAAGCTGCTGAGAACAGCG 3627
QY 4671 GATTAAGCTGCGGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 4730
Db 3628 GATTAAGCTGCGGCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3687
QY 4731 CCGAGGAGTATTCAGATCCAGCGGCTATGAGAGCGCTGTTTACGCTGAGGAGCAAGCA 4790
Db 3688 CTTACGCTTATTCAGATCCAGCGGCTATGAGAGCGCTGTTTACGCTGAGGAGCAAGCA 3747
QY 4791 CTGAGTATTAATGCTGAGCTGATGCGGCTGAGCGGCTGATGCTGCTGCTGCTGCTG 4850
Db 3748 CTTAAGATTAATGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3807
QY 4851 GCGAGGAGCTGAGCTGAGCAACATCAAGGCGGAGCTGAGGAGGAGATGAGCACTTT 4910
Db 3808 GCGAGGAGCTTCAATTAAGATTAAGCAAGGCGGAGCTTGAAGCGGAGATTAAGCAAG 3867
QY 4911 AATCCAGGAGTGTGCGGCTTACGCGGCTTCTGCGCTTATTAAGGAGATGCTGCTGCT 4970
Db 3868 AATCCAGGAGTGTGCGGCTTATGAGCGCTTCTGAGCTTATGAGCTGCTGCTGCTGCT 3927
```

QY	4971	GAGGCTATCTGAACCTGGACAACCGTACTCTTTCTGCAATTCGGCACTGCTGCTGCC	5030
Db	3928	GAGGCTATCTGAAGCTCGGCACCTGTGCTTTTATCGTTCAAAACATGCTTCGGGCC	3987
QY	5031	GGGCAAGAGCTGGGTGGGCTGCAGGCCCTGGACGAGTGTCAACGACTGGTGGCAATTT	5090
Db	3988	GGGCCCCAGTTTGGCAAGCTTGCAGGCATGGCGCAGCTTTACAGTAAAGTTGCCGGCTAC	4047
QY	5091	ATTCAATGACATCATCTGCGCCAGGCTTTTTCCTTTCACTCGTTGCTGGTCCGCGGTAT	5150
Db	4048	ATTGAGATGAGCATCTTCGGCAGGCGCTTTTCTTTTCACTCGCTCTTAAATGAGGGGAAT	4107
QY	5151	CCTTTTGAACAGCTCTTGATCTAATACCTTAAATCAAGCACTGGAGGCGAAATGGGCGAT	5210
Db	4108	CCGTTTGAACCTGTGCATTTTAAACCTGATATTCACCGTTAAGACGGAAATGGGCGTTC	4167
QY	5211	TGTTTTCGCGCGGCGGTACCGGTCGCGTGGTGGATGGACATGGCGCGGCTTTTCGCAT	5270
Db	4168	TGTTTTCCACCGGTTGAAACCGGTGGCTGGTCAATGGCATGATCAAGCTGTTTCAAGAT	4227
QY	5271	TTGGCGGTTGAACCTGCTCTCAACGCCGAAGTCAGCCAGCTGGAGACCGAGGTTAACCCG	5330
Db	4228	CTGGGCGCGGAAAGTCGTGCTTAAAGCCCGGCTCAGTATATGAACACCGTTGGGAAACAG	4287
QY	5331	ATCAGCGGTGTCCAGCTGAAGATGGGCGCGCTTTTGCCTCCCGCGCGCTTGTGCTCAAT	5390
Db	4288	ATTCAAGCCGTGTCAAGTTGAGAGACGGCAGACGGTTTAAACCTGCGCGGTGGCTGAAAC	4347
QY	5391	GCTGACGTGTGCATACCTACGATCGCTGTTAAGCAAGCATCTGTGGGGCGTTAAACGC	5450
Db	4348	GCTGATGTTGTGATACATCCCTATTCGCGATCTGCTGTCTCAGACATCCCGAGCGCTAAGCAG	4407
QY	5451	GCGGCAAGCTGGAAGCGCAGCGGATGAGCAACTCGCTGTTTGAATCTATTTTGGCTCTT	5510
Db	4408	GCGAAAAACCTGCATTCAGACGATGTAGTAACTCACTGTTTGTACTCTAATTTTGGTCTC	4467
QY	5511	AATCATGCCCAACCCGACGCTGGCGCACCAACGAGTGTGCTTTGGTCCGCGTATCGTGA	5570
Db	4468	AACATCATCAACGATCACTCGCCCATCATACCGTCTGTTTGGGCCACGCTACCGTGA	4527
QY	5571	TTGATTCGATGAGATCTTCAATATAGCAACGAGCTGGCGGAAAGTTTCTGCGTATCTGCAT	5630
Db	4528	CTGATTCAGAAATTTTAAACGATGAGTGTCTGGGTGAGGATTTTTCGCTTATTTTACAC	4587
QY	5631	GGCCCTGCTCCAGCGATCCGTGCTGGCACCGCGGGCTTCGCGCAGTTTTCAGTCTG	5690
Db	4588	GCACTTGTGTCAACGATTCCTGTCACTGGCACCGGAAGGTTGCGGACGATTAATGTCTG	4647
QY	5691	GCGCCGGTCCGCACTCGGTATCCGCCCAATTTGACTTGGCAACAGAAAGGCCCGCGCTG	5750
Db	4648	GGCGCTGTCCACACTTGAAGGCACGGCAACTTCGATGGGGCGTGAAGGACCCCGACTG	4707
QY	5751	CGCGATCGATTTTTCCTATCTGAGAGGACCTATATGCGGGTCTTGGCAACAGATTA	5810
Db	4708	CGGATCGATTTTGAATCTACTTATAGCAACATTAATGCTGTGGCTTGCAGAGCAATGG	4767
QY	5811	GTTGACACACCGTATGTTTACGCCGTTTGAATTTTGCAGACAGCTGACCGCATCAGGCC	5870
Db	4768	GTTGACACACCGTATGTTTACGCCGTTTGAATTTTGCAGACAGCTTCAAGCTTGCAGAGT	4827
QY	5871	TGAGCTTTTCCGTGGAACCCATTTTGAACGAAAGCGCTGGTTCCGCGCGCATTAACCGC	5930
Db	4828	TGCGCCCTTCGCTTGAACCTATCTGACCCAGAGCGCTGTGTTCCAGCACTAACCCG	4887
QY	5931	GATGCGGATTAATTAACCTTTATCTGTGGGGGCTGGCACGCAATCCCGGTGCGGTGTG	5990
Db	4888	GATTAAGCATTTGATTAATCTTTATCTGTTTGGCGCAGGCAACCAATCTGGCGCGGCAATT	4947
QY	5991	CCAGGCGTATCGGCTCCCGGAAAGACACCGCCAGCTGATGTGTGGAAGATCTGACCGGA	6050
Db	4948	CCGCGGTATCGGCTCCCGGAAAGGAGGAGCGGACGCTTAATCTGTGAAGACCTGATTTTGA	5007

QY	6051	TGAACCAACCCGCGCTGATTGACAGAGTACGCAACCAATGGCGCAGGGCTCCAAAAGT	6110
Db	5008	CGAATA---CGTATTACTGAATCATGCGCGTCGAAACCAATGGCGGTTGGCTGAAAAGCT	5064
QY	6111	TCGCAACGCGTACCCGGCTATTGGATCTTTCAACGCGCGCAGTACGCTGATGCTGTACG	6170
Db	5065	TTGCGACTGCATGACGCTTTTGACGCGCAAAACCCGTGCAAGCGTGTGATGCTTTAAG	5124
QY	6171	CCTGGTGTGCTCACTGTGACGATGTGTATGATGTGTACAGACGCTGGCGAAAGCGGACGC	6230
Db	5125	CATGTGCGCGCACTGGCGAGCAGTCACTGACATCAACACTGCGGCTTTTCATGCGGACC	5184
QY	6231	AGCACGCGGTTGGCGGATGACACAGCGCGGATGTGGCCACTGCAATTCGAAACCCGCGCGC	6290
Db	5185	AGCCCTCTTCGMAATGCTGTAGCAGCGCCTTCGACAGACTTTGAAAATGAAAACCCGTCAAG	5244
QY	6291	CCTACAGCGGTGCCCATGATGATGAACACAGCGTTTGTGTCTTTCAAGAAATGTGGCTTA	6350
Db	5245	CCTAGCGCGTTGGCAATGACAGAGCCGCTTTTTCGCGGTTTCAGAGGTGTGCGCATATG	5304
QY	6351	CGCATCAGCTTCCCCAGACGTGCTTTTATATCATGTGAAGGGTTTGGATGGATGGCG	6410
Db	5305	CGCATGATATGCTTCCGCGCTTACGCGCTTGCACACTGTGAAGGTTTGTCCATGATGTGC	5364
QY	6411	GTGAAGAACGTTATGCTGTGTTTCGGGAGCACGCTGTGCTTACGCTATCAAGTGGCCGCGC	6470
Db	5365	GCGAAACCGGTACTCTGACACTGAGCAGATAGCGTGTGCTTATGCTATACGTCGCGGTTG	5424
QY	6471	TGTTGGGTTATATATATGGCGCGCGTATGTGGCGCTTACGTATGAGCGCGTACTGATACG	6530
Db	5425	TTGTGGGCGCTATATATGTGGCGCAATTTATGGCGCTTCCGATTAACGCCACGCTGATTCGC	5484
QY	6531	CCTGATATTTGGGCTGTGGCGTTTCAGGCTTACCAATATCGCACGGATATCGTTGAGACG	6590
Db	5485	CCTGGGATCTCGGGCTGGCTTTCCAGTTGACCAACATTTGGCGGTATATTTGACGATG	5544
QY	6591	CGAGAAATGGCCGTTGCTATTTGCCCAAAAGCTGTGTGATGAGGCCGGACTGAGCGCGC	6650
Db	5545	CTCAGATGGGCGCGTGTATCTGTGCTGAAGCTGTGGAAGGAAAGAAAGACTGACGAAAG	5604
QY	6651	CCGACCTTGCAGATCGGCAATCGCAGCGAGCGCTGGCCGCGTGGCAGCCGCTGTGGTGC	6710
Db	5605	CGAATTAATGCTCGCCAGAAAACGGCAGGCTTTAAACCGTATGCGGGCGACGTGTAC	5664
QY	6711	GCGAGGCGGAGCGCTACTATCAGTCAAGCGCGCAGCGGCTGTCCAGAAATTGCGCTCGCTT	6770
Db	5665	GGAAAGCGGAACCTTAATTACGATTCATCAATGAGCGCGCTGTGGCACAATTAACCTTAAGCT	5724
QY	6771	CGGCGTGGGCGATCGCCACCGCGCGCGCGCTTACCGGGAAATTGGGGTTAAAGTGACG	6830
Db	5725	CGGCGTGGGCGCAATTCGACAGCGAAGCAGGTGTACCGTAAATTTGGGTGAAAGTTGAAC	5784
QY	6831	ATGCGCGGTGCCGGGCATGAGGATACGCGCACGCGCACCAATGAAGCGAAAAGCTGTGCGC	6890
Db	5785	AAGCGCGTAAAGAGGCGCTGTGGATTCATGCGCAGTCCACGTCACCGCGAAAATTAACGC	5844
QY	6891	TGCTGTGAAGGTGCGCGCGCTCGGCTTACTTTCGCGCTTGTCTCATTCGAGGGCGGCTC	6950
Db	5845	TTTTGTGACGCGCATCCGTCAGGAGGTACTTCCCGGATGAAGAGTATCCACCCGTC	5904
QY	6951	CTGCGCGTGTGTGGCAGGCTCCGCTTTACACAGACGCCCATGTGGCGTGGCGCAGCGTGC	7010
Db	5905	CTGCTCATCTTCGGCAGGCGCCG-----ATCTAGCCGCAATGCCCTTCTCTCAGCGTGC	5958
QY	7011	CTGCACTTGTGCAACGCTGTGGGCGTTAAAGAAAGCAAGAAAGCAAGCCTTCCGCTCC	7070
Db	5959	CTGAAGTTAGATTAACGTGTGGCGCTTACAGAAAACAAAGGACACGCAAGCCCTCTTTTCC	6018
QY	7071	CGGCAACGATATGATGACGCTGTGGCGCAATATACACGCTTTCAGATAGCCTTTTGTGCTGG	7130
Db	6019	CTTTAAGATATATGATACGCTGGGCCATGTATTAACGTTTCAGGTAGCCTTTGCGCGG	6078
QY	7131	AATATGCGAAACGGCGCAGCGTTTATGACAGCGCGCTCATGACATATAATGAGAGC	7198



Db 6079 TAGTACGCGAAGCGCCAGCGCTGTGTATCAGCGCTCGTGAACCAATAAATACAGTAA 6138  
QY 7191 GCCATAGGTCGTATGCGACGCCCAATCCACTGACGGCGCAGACGCCGTTGACACCCAG 7250  
Db 6139 ACCATAGCGCGTATGCGCTGACCAATCCACTGAGGGCGCAGATTCTGTACTGCCGAA 6198  
QY 7251 CCAATACAGGCCAATCGACACAGCGCAACACCGGCGATACAGTGTGTTAGCTAAA 7310  
Db 6199 GTAAATCAGGGCAATCGACCAATGGGAAATACCAAGCGCATAGAGATCGTTAACTTCAA 6258  
QY 7311 TTGGCTCTCATGTGTGTCATGATGCGCAATGCGACGCCCATCCAGCCATGACATAT 7270  
Db 6259 TGGCGCTTATGCGGCTTATGATGTGAAGATGCGACCCCAACCCAGCGTGCATGAT 6318  
QY 7371 GATTTATGCGACAGCGCCGCGAGATCTTCATCATCATGATACCGGTACCGACGATPAG 7430  
Db 6319 GATTTATGCGACGATGACCAACCACTTCATGCGCGACCGGTGACAAACAGATCAG 6378  
QY 7431 CGCATTCATACCAAGCATCATTTGTCATTTGCGAAGAGTGAAGTAAAGTGA 7490  
Db 6379 GGCATTCGAATCACAACATATCTCTCCAGTACAGACGTCGGGATCAAAAGTGA 6438

RESULT 3  
ERECT 6918 bp DNA linear BCT 24-AUG-2002  
LOCUS Erwinia uredoovora DNA carotenoid biosynthesis gene cluster (crte,  
DEFINITION crtx, crty, crtl, crtb, crtZ), complete cds.  
ACCESSION D90087  
KEYWORDS D90087.2 GI:22474502  
SOURCE Pantoea ananatis  
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
1 (bases 1 to 6918)  
REFERENCE 1 Nakamura, N., Nakagawa, M., Kobayashi, K., Yamano, S., Izawa, Y.,  
AUTHORS Nakamura, K. and Harashima, K.  
TITLE Elucidation of the Erwinia uredoovora carotenoid biosynthetic  
pathway by functional analysis of gene products expressed in  
Escherichia coli  
J. Bacteriol. 172 (12), 6704-6712 (1990)  
JOURNAL 91072214  
MEDLINE 2254247  
PUBMED On Aug 23, 2002 this sequence version replaced gi:216681.  
COMMENT These data kindly submitted in computer readable form by: Norihiko  
Misawa  
Marine Biotechnology Institute  
3-75-1, Heita, Kamaitshi-shi,  
Iwate 026-0001  
Japan  
e-mail: norihiko.misawa@mbio.jp  
Phone: 81-193-26-6581  
Fax: 81-193-26-6584

FEATURES  
source location/Qualifiers  
1. .6918  
/organism="Pantoea ananatis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:553"  
/note="carotenoid biosynthesis genes"  
225. .1133  
/gene="crte"  
225. .1133  
/gene="crtE"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAAL1424.1"  
/db\_xref="GI:216682"  
/translation="MTVCAKRVHLLTRDAEQULADIRRLDQLPEVGERDVVGAM  
REGALAPKRIKRIEMLLTPARDGCAVSHGDLACAVEMVAASLIIDPMCMDDA  
KLRGRPTIHSYGEHVAITAAVALSKAGVADGLPLAKRNVASLTAIIMO  
GLVQGGPKDLSBDKRSAAAILMTNHTFKSTLFCASMQMASTIVANAASSARCLHRF

gene 1143. .2438  
CDS /gene="crtX"  
1143. .2438  
/gene="crtX"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAAL1425.1"  
/db\_xref="GI:216683"  
/translation="MSHPAALAPFYSYHVALONLAQELVARGHRTYFIQYDIKHLI  
DSEITFHSYDSDHPGALATRYLHLAHPISMLKINEMARTDMLCRELPQAN  
DLAVGVIVDQKBPAGALVABNLGLPISVACALPINRBDMLAVMPEYGSADAR  
ERYABEKIYDMLMRHNDVIAHSHRMGLAPQKIHQCFSPLAQISQLVPEIDPRK  
ALPACFHAVGPLETHAPSTSSRYFTSSKEPIFASLGLQGRYGLFETIVACEE  
IDQGLAHAGRLTDCBELASRHTQVDPADQSAALSOQLATHGGMVNLVDAI  
NRTPLALPLAFEDQGVASRIYVHIGIKRASFTSHALARMRSLLTNVDQQRMA  
KIOTLRLAGTMAADIIEQWCTGQPVLLSGSVATL"  
2422. .3570  
/gene="crtY"  
2422. .3570  
/gene="crtY"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAAL1426.1"  
/db\_xref="GI:216684"  
/translation="MQPHYDLILVAGLNGLIATLRLQOQPDMLILLIDAPQAGN  
HTWSFPHNDLTESQHRMIAPLVVHMPDQVPRPTRRKLNISGFCITISQRFVILOR  
OFQPHLMPTAVAEVNAESVRLKKGQVIGARAVIDGRVANAALSVGFQFGEWR  
LSHPGLSSPILMDATVDQONGRFVYSLPSPTRLIEDTHYIDNATLDEPQRONI  
CYRAOQGMQLOTLREBOGALPITLSGNADAQWQORPLACSLRGLPHPTTGYSP  
LAVAVADRSLADVFTSASIHAIHIFABRWQOQCFRMLNMLTLACGADSRMRM  
QRFYGLPEDLIAFYAGKLLTDLRLILSGKEPVVLAALQALMTTRH"  
3582. .5060  
/gene="crtI"  
3582. .5060  
/gene="crtI"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAAL1427.1"  
/db\_xref="GI:216685"  
/translation="MKPTTVIGGFGGLALATRLQANGLPVLLLEQRDKRGRAVYE  
DQGFPTDAPRYITDPSAIEELFALAGKOLKEVELLPPTPRFLRMEGKQVNVND  
QTRLEAQIOQFNEDVEGYRQFLUDISRAVFKESYGLGVPPLISFEDMLBAPQAKL  
QAMRSVYSVASYIEDLRLQAFSHSLVGGVGFATSSLYTLIHLERWGVMPFPG  
GHTYRDLSEHPAAVQSNKLOTKRNNSLFTVLYFGLNHHLDQALHTVCFGRYEL  
IDEIFNHDGLADBSLYLHAPCVTDSLAPEGGSYVLAAPVPHLGTANLDMVEGPK  
LADRIYAYLEQHTMPELSQULYTHRFTFPDRDQANAHGSAFSVEPVLTQSMRFP  
HNDDKITNLVVGAGTHPGAGIPVIGSAKATAGLMDLI"  
5057. .5986  
/gene="crtB"  
5057. .5986  
/gene="crtB"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAAL1428.2"  
/db\_xref="GI:22474503"  
/translation="MNPSLNLNVAETMAVGSKFATASKLPDAKTRRSVLMYAMCR  
HDDVDVDDQTLGROAROPALOTPEOILNKEKTRVAVGSGMHEBAPAFAPQVAMAH  
DIAPAYFPHLSGFAMDVAREAOQSQDLTLRTCYHAGYVGLMADIMKGRDATTDR  
ACDLGLAFQULNARDIVDAHAGRCYLPASWIEHGLNKENYAAAPENQALSRITAR  
LVQEARPYLISATAGLAGPLRSAMAIATAKOVYRKIGKVEQAGQADQDROSTTTP  
EKTLTLLAASQALTSRMRAPPRPAMLMORPL"  
complement(5925. .6452)  
/gene="crtZ"  
complement(5925. .6452)  
/gene="crtZ"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="BAAL1429.1"  
/db\_xref="GI:216687"  
/translation="MLMIMNALIVPVTYIGMEVIAALAIKTYIMHGKMGMLSHHEPR

## ORIGIN

KGAPEVNDLYAVVPALSTILIIYIGSTWMPLOWIGAGMTAYGLYPMNDGLHWOR  
PFRIRKGIKRLIMHMRHMHAVRKEGCVSFGFLYAPPLSKLOATLIRHRKARQA  
ARDAQGEDEBPASGR"

Query Match 30.2%; Score 2618; DB 1; Length 6918;  
Best Local Similarity 64.7%; Pred. No. 0; Mismatches 2165; Indels 24; Gaps 7;  
Matches 4011; Conservative 0;

QY 1284 CTTGAAACATTTATTTGCTGCGGAGAGGCGGATCGCGTTGCGGCGCGGATCGGCGC 1343  
DB 303 CTTGATCACTATTGCGCGTGGAGGAGAAAGGAGATGTTGGGTGCGCGATCGGAA 362  
QY 1344 GGCAAGCTGGGACCGGCGGAAAGTATTCCTGCTGCTGCTGCGAGCAAGGAT 1403  
DB 363 GGTGCGCTGGGACCGGCGGAAAGTATTCCTGCTGCTGCTGCGAGCAAGGAT 422  
QY 1404 ATGGGCTGTGAGTGGGCGAGAGGCGATCCTGATCTTGGCTGTGCGGATGATGATG 1463  
DB 423 CTGGGTTGCGCTGCTGAGCGATGACGATTAATGATTTGGCTGTGCGGATGAAATGTC 482  
QY 1464 CAGCGTGCCTCACTGATCTTGCAGCAATTCATCAATGATTAAGCCCGGATGCAAGT 1523  
DB 483 CAGCGGCTTGCCTGATCTTGCAGATGCGATGCGCTGATGAGCAATGCGAAGCTGCGCGC 542  
QY 1524 GGGCGCCCGGCAATCCATGCTGATATGAGGAAAGCTGCGATCTTGGCGAGCGGCGC 1583  
DB 543 GGAAGCCCTACCATTCATTCATTAAGAGAGATGAGCAATATCGGCGCGGCTGGCC 602  
QY 1584 CTACTCAGCGCGCTTGAAGTATGCGCTCGCGCGCGGCTGCGGAGCAAGCAACACAAA 1643  
DB 603 TTGCTGATTAAGCTTTGGGCTAATTCGATGCAATGCGCTCAAGCCGCTGGCAAAA 662  
QY 1644 GCCGAAGCATTTGCGAGCTCTCTGCGCGTGGCGCTGAGGAACTGCTTCAAGGTCAG 1703  
DB 663 AATCGGCGGTTTCTGAAGCTGTCAAAAGCATGCGCATGCAAGATTTGGTTCAGGTCAG 722  
QY 1704 TTCCAGATCTGATGAGCGGCGCACAGCGGAGTCCGGAAGCATCAAGCTGACCAAT 1763  
DB 723 TTCAAGATCTGTCTGAAGGAGATGAGCGGCGAGCTGAAGCTATTTGATGACAT 782  
QY 1764 GAATGAAACCAAGGCTGCTGTTGCGCGCAAGCTGAGATGCGGCGAT 1823  
DB 783 CACTTAAACCAAGAGCTGTTTGTGCTCAATGAGATGCGCTGATTTGCGGAT 842  
QY 1824 GCGTCAAGTCAAGTCAAGCTGTTAAGTATTTGCGGAGATTTAGTCAAGCTTTC 1883  
DB 843 GCGTCAAGGAGCGGCTGATTTGCTGATGCTTTTCACTTGAATTTGTCAGGCAATTT 902  
QY 1884 CAGTTACGAGAGAGCTGCGGAGATGCTTAAGCAACCGGCAAGGATGCTGATCAGGAT 1943  
DB 903 CAACTGCTGAGATTTACGATGAGTACGACCGGATGAGGATGAGATCAAGGAC 962  
QY 1944 CAGGCAATCAAGCTGCTGAGATGCTGAGGCGCGGAAGGCGCTGAGCGTCTGCGC 2003  
DB 963 GCCGTAATGAGCGCTGCTGATCTGTTAGGCCGGAAGGCGTTGAAGATGCTGAGA 1022  
QY 2004 GACCATCTAAGAGCGCGATGACACTTGGCTGCGCTGCTGATCGGCGTGTGCGCAC 2063  
DB 1023 CAACATCTTCAAGCTTGCAGTGAAGCATCTCTGCGGCTGCGACACAGGCGACGCGCACT 1082  
QY 2064 CGTCAATATGAGAGCGGCTGTTTAATCAAGAGTGGAGATGTTCAATGAGAGCGGCA 2123  
DB 1083 CAACATTTATTTCAAGCGCTGTTTACAAAACCTGCTGCGGCTGATTAAGATGCT-- 1140  
QY 2124 TACCTATGAGGATTTGCGCTTATTTGCGCAAGCGCTCAAGCACTTTCAAGCATTTGC 2183  
DB 1141 ---GATATGACATTTGCGGAGATGCGACCGCTTTTAAAGGCAATGTCGCGCATTAAC 1197  
QY 2184 AGGCGCTGCGCAAGCGCTGCGGCGGCGGACATGCACTTATTCAGCAAGATG 2243  
DB 1198 AGAATCTGCTCAGAACTGGTGGCGCGGCTGATGCGGATGACCTTTATTCAGCAATACG 1257

QY 2244 ATGACAGCACTTGTGAGGACAGAGCGCATTTGTCGCGCTGCGGAGCGGACGCG 2303  
DB 1258 ATATTAACAATTGATGATAGGAAACCAATTGATTTCAATTCCTGCGGACAGACAGCC 1317  
QY 2304 ATCTGCGGATGCTCTCCAGGAACTCAGGCGCTGCGGACCGGCGGCTGTCG 2363  
DB 1318 ATCCCCCGCGGCTTAACGCGGCTGCTACACTGCGGCTGATCTCTGCGGCGCGTCA 1377  
QY 2364 TGTTCGCGGATTCAGATCTGCGGACAGCAACGATATGATGCGGAGACTGCGCG 2423  
DB 1378 TGCTGAAGCTCAATGAATATGCGGCGGACCAAGATATGCTGCGGAACTCCCC 1437  
QY 2424 CGGTGCTCAACGAGCTGAGGCTGATGCGCTGATGCTGATCAATATGAAGCGGCTGCTG 2483  
DB 1438 AGGCAATTAAGATCTGCGGCTGATGCGCTCATTTGATTAAGATGAAGCGGAGGCG 1497  
QY 2484 GTCTGTGAGAGGCGTTACAGCTGCGCTTGTGTCGCTGCGCTGCGGCTGCGCTCA 2543  
DB 1498 CGCTGCTGAGAGCACTGGAAGCTGCGCTTATCTGTCGCTGCGGCTGCGCTCTCA 1557  
QY 2544 ATGCGAAGGCGCATTCGCGCTGCTGATGCGCTTGTGCTGCTGCTGCTGCTGCTCA 2603  
DB 1558 ATGTAACCGGATATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTCA 1617  
QY 2604 CGCTGAGCGCTATCAGGCGGACGAGCATCTAAGACCGCATATGCTGCTGATGCG 2663  
DB 1618 CTGCGAAGCTTATGCGGCGGAAATTTATGATGCTGCTAATGCGTCTGATGACC 1677  
QY 2664 CTGCTATGCTGCTGATGCGGCGCTTGGCTGCGGCGGCGGATGCTTATCATCAT 2723  
DB 1678 GTGCTATGCTGCAACAGGCAAGATGCGCTTGAAGCTTCAAGCTTCAAGT 1737  
QY 2724 GTCTGCGCGCTGCGGCAATGATGATGCTGCTGCTTGTGATTTTCAAGCGGACG 2783  
DB 1738 GTTTTTCGCACTGCGGCAATGATGATGCTTGTGCTGATGATTTTCCCGGAAAG 1797  
QY 2784 AACTGCAAGCTGCTATCAGAGCGTGGTCCGTCGCG--GATCTCAAGTGTAGCGGCG 2840  
DB 1798 GCTTACCGGCTTGTTCATGCGGCTGCGGCGGCGGCAAGCAAGCGCTCAAGT 1857  
QY 2841 CGCTGCGGCAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2900  
DB 1858 CTTCATCCGTTATTTTAACTCTCAGAAACCCCGGATTTTTCGCTGCGGCGGCGG 1917  
QY 2901 TACAGGCGCATGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2960  
DB 1918 TTCAAGGACACCTTATGAGGCTGTTTAAACGATGAGAAAGCTGTGAAGAAATGACG 1977  
QY 2961 TGTGCTGATGAGCAACTGTGCGGCTTGAAGCGGCGGCGGCGGCGGCGGCGGCGG 3020  
DB 1978 GTCAAGCTCTGTTAGCCCACTGAGTGTGCTTACGAGCTCTCAAGTGAAGAGCTGCGC 2037  
QY 3021 TGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3080  
DB 2038 GAAAGCGCTATACAGAGTGTGATTTTTCAGATGATGAGCGGCTGCTGCTGCTGCTG 2097  
QY 3081 AACTGTTATCACTCAGCGGCTGTAACAGTGTGCTGTAAGCACTGAGTGTGCGGCG 3140  
DB 2098 AGCTGAGATACCAAGGCGGCGGATGATGATGATGATGATGATGATGATGATGATG 2157  
QY 3141 CGATGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 3200  
DB 2158 CCGTTTAAAGCGCTTCCGCTGCGCTTGTATGAGCGCGGCGGCTGCGGCTGCGGCTGCG 2217  
QY 3201 AGCGGCTGCGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 3260  
DB 2218 ACAGCATGCGCAAGCGGCTTCCGCTTATCAACAGGCAAGCTTGTGCTGCTGATGATG 2277  
QY 3261 AACAGTTGCTGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3320  
DB 2278 GTTATTTGCTGACCAAGCTGATCTTTTCAAGAGCGGCGGAAATCAAGACAGCCCTTC 2337  
QY 3321 AGCGGCGGCTGCTGATGCGGCGGCGGCTGATTTGTGAGCAAGCGGCTGCTGCTGCAAGC 3380

Db 2338 GTTGGAGGGGACACATGCGCGCTGCCATATCATAGCAGGTTATGTG--CACCGGT 2396  
Qy 3381 AAATCGCTGCGGAGGCCACTGATGCGCGCACTTATGATGTCACTTGTGCGGTG 3440  
Db 2397 CACCGCTGCTTAATGTGGAGCGCGCTATGCAACCGCATTTATGATCTGATCTCGTGGGGG 2456  
Qy 3441 CGGCGTGGCTAAGGGGCTGATTTGCGCTGCGTTTACCGCAGCTCAGGCCGCCACTTAAGT 3500  
Db 2457 TGGACTGCGAATGGCCTTATCGCCCTGCGCTTTCAGCAGCAGCACTATATGCGTAT 2516  
Qy 3501 TTGCTACTGAGAGTCAAGCGCGAGCGCGCAATCATACCTGCTGCTTCAACGCGA 3560  
Db 2517 TTGCTTATGAGCGCGCAACCCAGGCGGGCGGGAATCATACGTGCTATTTCAACGCA 2576  
Qy 3561 AGACGTCAAGGAAGCGCATTTGCTGCTGCGAGCGCTGCTTTGCGCGCGCTGGCCGG 3620  
Db 2577 TGAATTTGACTGAGAGCCACATGCTTGGATAGCTCCGCTGCTGATCATCATGGCCGA 2636  
Qy 3621 TTATCAAGTATGCTTCCCAACCTGCGCTGCCAGCTGGAGTGTGAATATGCTCGATTG 3680  
Db 2637 CTATCAAGTATGCTTCCCAACGCGCTGCTGAAGCTGAACAGCGGCTACTTTGTATTAC 2696  
Qy 3681 CTGCGAGATTTTGGCCGCACTTATCAGCAGTGTGCTGCGCGCTACGCAACCGCAG 3740  
Db 2697 TTCTCAGCGTTTGGCTGAGGTTTACAGCAGCAGTTTGGCCCGCATTTGTGATGATAC 2756  
Qy 3741 GCCGCTCAGAGAGTCTCACCCACCGGGGTCAAGCTGGCGAGTGGCGGATGTTACAG 3800  
Db 2757 CGCGGTGCGAGAGGTTAATGCGGAATCTGTTGCTGTAATAAAGGATCAGGTTATCGGTG 2816  
Qy 3801 GCAGGGGTGATTTGAGGAGAGCGGGGCTGACCGCAGCAGCATCTGAGCTCGCTATCA 3860  
Db 2817 CGCGGTGATTTGAGGAGAGCGGGGCTTATGCGCAATTTCACTGAGCGCTGAGCTTCCA 2876  
Qy 3861 GGCAATTTGCTGCTCAGAGTGGCACTGCGCGCGCATGGCTGAGAGCCATATTT 3920  
Db 2877 GGGGTTATTTGGCCAGGAATGGCGATTGAGCACCGCAGTGTATTCGTTCTCCATTAT 2936  
Qy 3921 GATGACGCCAGCGTGCATCAGCAGCAGGTTATCGCTTGTATTACACCTTGCCTCAG 3980  
Db 2937 CATGATGCGCAGCGTGCATCAGCAAAATGTTATCGTTATCGTTACAGCTTGCCTC 2996  
Qy 3981 TGGCAGCGTTTACTGATTAAGATACCACTACATCAACCATGCGCAGCTGATGCGC 4040  
Db 2997 GCGGACCAATGTTATTTGAAGACACGCACTATATTGATATGCGCATTAGATCTGA 3056  
Qy 4041 ACAGCGCGCGCTCACATTACGATTAATGCGCACACGCGCGCGTGAATTTGGCGCAGT 4100  
Db 3057 ATGCGCGCGCAAAATTTTGGCATATGCGCGCAACAGGTTGGCAGCTTACAGACAT 3116  
Qy 4101 GCTGCGCAGAGCAGCGCTGCTGCGCATCAGCTCAGCGCGATATCATGATGTTCTG 4160  
Db 3117 GCTGCGAAGAAACAGGGGCGCTTACCCATTACTGTGCGGAAATGCGCACCATTTCTG 3176  
Qy 4161 GCAACGACGACAGCGGCAACCTGTCAGCGGGCTGCGCGCGGACTGTTTCAACCCAC 4220  
Db 3177 G---CAGCAGCGCGCCCTGCGCTGTAGTGTATACGTGCGGTCTGTTCATCTTCAAC 3233  
Qy 4221 CGGTTACGCTGCGCGCGCGCGGTGGCGTGGCGGAGAAATTGCGACAGCAGCTGCGCG 4280  
Db 3234 CGGCTATTTACATGCGCGCTGCGGTGGCGTGGCGACCGCTGATGTCATTGATGTTT 3293  
Qy 4281 CGACGCTCACACGCTGAGCAGCATGCAATCTTTGCGCCGTCAAGCATTTGGCGAGCA 4340  
Db 3294 TAGCTGGGCTCAATTTACCATTCATTCAGCATTTTGGCGGAGGCGTGGGAGAGCA 3353  
Qy 4341 GCGCTTTTTCGCTGTGTTAATGCAATGCTGTTTCTTGGGAGCGGCTGAAACAGCGCTG 4400  
Db 3354 GGGCTTTTTCGATCTGATGCAATGCTGTTTATGCGGAGCCGCGCATTTCAAGCTG 3413  
Qy 4401 GCGGTAATGAGCGTTTATACGGGCTTACGCGGATTTGATTAAGCGCTTTTACCGCG 4460

Db 3414 GCGGTTATGACAGGTTTTTATGTTTACCTGAAGATTTAATTTGCCGTTTTTATGCGG 3473  
Qy 4461 GCACTGCGCTTACGAGATTAAGCAGCATTTCTGTGGGCAACCGCGGTCCCTCTCG 4520  
Db 3474 AAAATCAGCTGACCGATCGGCTACGTATTTGAGCGGCAACCGCGCTGTTCCGTAT 3533  
Qy 4521 CGAAGGCGTGGGCAATTAATGATGACCTCTCGGTTACGAGGGAAGAAATTAAGAAC 4580  
Db 3534 AGCAGCATTTGCAAGCAATTAATGACCTCATCTTAAGGCGACTA---CATGAACCA 3590  
Qy 4581 ACCTATGATTTGCGCAGGCTTCCGTTGCGCTGCGCTGCGATTTGCTGCAAGCGCC 4640  
Db 3591 ACTAGGTTATTTGTCAGAGCTTCCGTTGGCTGGCACTGGCAATTTGCTACAGGCTGG 3650  
Qy 4641 GCGCTGCGGCTCAGCTGCTGGAACAGCGGATTAAGCTGGCGGGCGGCTTATGTTAT 4700  
Db 3651 GGGATCCCGCTTTCATCTGCTTGAACAGTGAATAAACCGCGCGGTGCGGCTTATGCTAC 3710  
Qy 4701 CAGGATCAGGTTTACCTTTGATGCGCGTCCGAGCGGTATTAACGATCCAGCGCTATC 4760  
Db 3711 GAGGATCAGGTTTACCTTTGATGAGCGCGGCGAGGTTATCAACGATCCAGTGCATT 3770  
Qy 4761 GAGCGCTGTTTACGCTGCGAGGCAAGCAACTCAGTATTAATGCTGATGATGCGGTG 4820  
Db 3771 GAGGATCTGTTTACGCTGCGAGGAAACAGTTAAAGATATGCTGAACCTGCTCGGTT 3830  
Qy 4821 AGCCATTTTATGCTGCTGAGGAAAGAGCGGAGCAGCTGAGCTACGAACAATCAG 4880  
Db 3831 AGCGCTTTTACCGCTGCTGTTGAGGAGTCAGGGAAGTCTTAAATTAACGATACAA 3890  
Qy 4881 GCGCAGCTGAGAGAGAGATGCACTTTTATTTCCAGATGTCGCGGTTACCGCCAG 4940  
Db 3891 ACCCGCTGAGAGCGGATTCAGCAGTTTAAATCCCGGATTCAGAGGTTATGCTCAG 3950  
Qy 4941 TTTCTGAGCTTATTCAGAGATGTTTCTGAGAGGCTATCTGAATCTGGCAACGTA 5000  
Db 3951 TTTCTGAGCTTATTCAGCGCGGCTGTTTAAAGGCTATTAAGTCTGCTGCTGCT 4010  
Qy 5001 TTTCTGATTTTCCGCGACATGCTGCGTGGCGGCAACAGCTGGGTGCGCTGACGCGCTG 5060  
Db 4011 TTTTATTCGTTAGAGACATGCTTCCGCGGCACTCAACCTGGCAATGCGAGCATGG 4070  
Qy 5061 CGCAGTGTCAAGCATGCTGCGGAAATTTATTCATGAGATATCTGGCGCAGGCTTT 5120  
Db 4071 AGAAGGTTTACAGTAAAGTGTGCGATGATCAATGAAGATGATCGCGCGGCTTT 4130  
Qy 5121 TCCCTTACCTCGTTGCTGCGCGGCTTATCTTTGCAACGCTTGATCTATACCTTA 5180  
Db 4131 TCTTTCACTCGCTGTTGGTGGCGGCAATCCCTTGCACCTCATCTTATATGCTTG 4190  
Qy 5181 ATTCAAGCATGAGCGCAATGGGCGTGTGTTTCCGCGCGCGGTACCGGTGCGCTG 5240  
Db 4191 ATACAGCGCGTGAAGGTGATGGGCGTCTGGTTTCCGCTGGCGGCAACCGCGCATTA 4250  
Qy 5241 GTTATGAGTGGCGCGGCTGTTTCCGATTTTGGCGGTGAATGCTGCTCAACCGCGAA 5300  
Db 4251 GTTCAGGGATGATTAAGCTGTTTCAAGATCTGGGTGGGAAAGTCTGTATTAACGCA 4310  
Qy 5301 GTACAGCAGCTGAGACGAGGGTTAACCGATCAGCGGTGTCAGCTGAAGATGGGCG 5360  
Db 4311 GTACAGCATTAAGAAACAGAAACAGAAATTAAGACCGGTGATTTAAGGCGGTGCG 4370  
Qy 5361 CGTTTTCGCGCGCGCGCTGCTGCTCAATGCTGACGTTGATATCAATGCTGCTG 5420  
Db 4371 AGGTTCTGACGGAAGCGTCCGCTCAATTTGAGATGCTGCTATCATCTATGCGACCTG 4430  
Qy 5421 TTAAGCAGCATCTGCGCGCGCTTAAACGCGGCAACCTGAAGCGCAAGGATGAC 5480  
Db 4431 TTAAGCAGCAACCTGCGCGGTTAAGCAGTCAACCAACTCAAGCTAAGGCGATGAGT 4490  
Qy 5481 AACTGCTGTTTGAATCTAATTTGCTTATTCATGACCCAGCGGAGCTGGCGACCA 5540  
Db 4491 AACTCTGTTTGTGCTTATTTGTTGTTAATCAACATCATGATCATGCTGCGGCTAC 4550

QY 5541 ACCGTCCTTTGGTCCGCGCTATCGTGAATTGATGATGAGATCTTCAATAGACCCAG 5600  
DB 4551 ACCGTTTGTTCGGCCCGCGCTTACCGGAGCTGATTAACGAATTTTAAATCATGATGTC 4610  
QY 5601 CTGGCGGAAGATTCTCGCTGATCTGCATGCGCCCTGCTCCAGCAGATCCGTCCGTGGCA 5660  
DB 4611 CTGGCAGAGACTTCTCATCTTATCTGCACGCGCCCTGTGTCAAGATTGCTCATCGCG 4670  
QY 5661 CCGGCGGAGCTGCGGACAGTTTTCAGTCTGCGCGCGGTGCGCATCTCGGATCCGCCGA 5720  
DB 4671 CTTGAAGTTGCGGCACTTACTATGTTGGCGCGCGTGCAGCTTTAGCAGCCGCAAC 4730  
QY 5721 ATTGACCTGGCAAGAAAGGCGCGCGCTGTGCGCATGGAATTTTGTCTTATGAGAGAG 5780  
DB 4731 CTCGACTGAGACGTTGAGAGGCGCAAACTAGCGACGATTTTGTGGTACCTTGACGAG 4790  
QY 5781 CACTATATGCGGAGCTGCGCAGCAGAGTTAGTGAACACACCGTATGTTTACGCGCTTGTAT 5840  
DB 4791 CATTACATGCTGCGCTTACCGAGTCAAGTGTGACGACCGGATGTTTACCGCGTTGAT 4850  
QY 5841 TTTGCGACACAGCTGCAAGCGCATCAAGGCTCAGCGTTTTCGCTCGAACCCATTTTGAAG 5900  
DB 4851 TTTGCGACACAGCTTAAATGCTTATCATGCTCAGCGCTTTCGTGAGCGCGTCTTAC 4910  
QY 5901 CAAAGCGCTGCTTCCGCGCGCATACCGGATGCGCACTTACTTAACCTTTATCTGAGT 5960  
DB 4911 CAGAGCGCTGCTTCCGCGCGCATACCGGATGCGCACTTACTTAACCTTTATCTGAGT 4970  
QY 5961 GGGGCTGCGACGATCCCGGCTGCGGCTGCGCAGCGTATGCGCTCGCGCAAAAGCGAC 6020  
DB 4971 GGGGCTGCGACGATCCCGGCTGCGGCTGCGCAGCGTATGCGCTCGCGCAAAAGCGAC 5030  
QY 6021 GCCGAGCTGATGTTGAGAGATCTGAACGAGTGAACAAACCGCGCTGATTTGAGAGTCA 6080  
DB 5031 GCAGGTTGATGCTGAGAGATCTGAACGAGTGAACAAACCGCGCTGATTTGAGAGTCA 5087  
QY 6081 CGCAACCATGCGCGCGCGCTTCCAAAGTTTCCGCGCGCATACCGGCTATTTGATCCTT 6140  
DB 5088 TCGAAGCATGCGCGCGCGCTTCCAAAGTTTCCGCGCGCATACCGGCTATTTGATGATA 5147  
QY 6141 CAACGCGCGCGAGTACGCTGATGCTGATACGCTGATGCTGATGCTGATGATGATGATG 6200  
DB 5148 AAACCGCGCGCGAGTACGCTGATGCTGATGCTGATGCTGATGCTGATGATGATGATG 5207  
QY 6201 ATGCTCAGACGCTGAGGAGAGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 6260  
DB 5208 ACAGTACAGCCTGCGGCTTTCAGGCGCGCGCAGCCTGCTTACAAACCGCGCAACAGCTC 5267  
QY 6261 TGGCGCACCTGCAATATGGAACCGCGCGCGCTTACAGCGGTCAGCGGTCAGATGATGATG 6320  
DB 5268 TGATGCAATCTTGAATGAAACCGCGCGCGCTTACAGCGGTCAGCGGTCAGATGATGATG 5327  
QY 6321 CGTTTCGCTCTTTCAGGAGTGGCGCTGAGCAGTCAAGCTTCCCAAGCAGCTGCTTTT 6380  
DB 5328 CGTTTCGCTCTTTCAGGAGTGGCGCTGAGCAGTCAAGCTTCCCAAGCAGCTGCTTTT 5387  
QY 6381 ATCATCTGGAAGGCTTTCAGGAGTGGCGCTGAGCAGTCAAGCTTCCCAAGCAGCTG 6440  
DB 5388 ATCATCTGGAAGGCTTTCAGGAGTGGCGCTGAGCAGTCAAGCTTCCCAAGCAGCTG 5447  
QY 6441 CGCTGCGCTTACGCTATCAGGTCAGCGCGCTGAGGCTTATGATGAGCGCGCGTATGAG 6500  
DB 5448 CGCTGCGCTTATGCTATCAGGTCAGCGCGCTTATGAGGCTTATGATGAGCGCGGATGATG 5507  
QY 6501 GCGTACGATGAGCGCGCTATCAGTCAAGCTTATGATGAGGCTTATGAGGCTTATGAGCT 6560  
DB 5508 GCGTACGATGAGCGCGCTATCAGTCAAGCTTATGATGAGGCTTATGAGGCTTATGAGCT 5567  
QY 6561 CCAATATGCAAGGAGTATGTTGAGAGCGCGGAGATGAGCGGCTTATGAGGCTTATGAG 6620  
DB 5568 CCAATATGCTGCGATATGTTGAGAGCGCGGAGATGAGCGGCTTATGAGGCTTATGAG 5627

QY 6621 GCTGCTGATGAGCGCGAGCTGAGCGCGCGCGAGCTTCCGATCCGCAATCGCGAG 6680  
DB 5628 GCTGCTGAGAGCTGAGAGGCTTGAACAAAGAGATTAATGCGGACCTGAAACCGTCAAG 5687  
QY 6681 CGCTGAGCGCTGAGAGCGCGCTTGTGAGCGAGGCGCGAGCGCTTCAATCAATCAAG 6740  
DB 5688 CGCTGAGCGCTTATGCGCGCGCTTGTGAGCGAGGAGAGAGAGAGAGAGAGAGAGAG 5747  
QY 6741 GCAAGCGGCTGAGAGATGAGCGCGCTTGTGAGCGAGGAGAGAGAGAGAGAGAGAGAG 6800  
DB 5748 CAGCGCGCTGAGAGAGGCTTGTGAGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5807  
QY 6801 TTTACCGGGAATTTGCGTAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6860  
DB 5808 TTTACCGGGAATTTGCGTAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5867  
QY 6861 AGCGCACAGTAAAGCGGAAAGCTGAGCGCTGCTGTTGAAGGTTGCGCGCTTCA 6920  
DB 5868 AGTCAACAGACAGCGCGGAAATTAACGCTGCTGCGCGCGCTTGTGAGAGAGAGAGAG 5927  
QY 6921 CTTTCGCGCTTGTCTCATCCGAGAGCGGCTGCGCGCTTGTGAGAGAGAGAGAGAGAGAG 6980  
DB 5928 CTTTCGCGAGTGGCGGCTCATCTTCCCGCGCGCTGCGAGCTTGTGAGAGAGAGAGAGAG 5981  
QY 5981 ACAGAGCGCATGAGCGCTGAGCGAGCGCTGAGCGCTTGTGAGAGAGAGAGAGAGAGAG 7040  
DB 5982 TCTAGCGCGCATGCTTCCCGAGAGCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6041  
QY 7041 AAAGCGCAAGAGAAACGAGAGCTTCCGCTCCCGCAGCGCATGATGATGAGGAGAGAGAG 7100  
DB 6042 GAAAGCGCAAGAGAAACGAGAGCTTCCGCTCCCGCAGCGCATGATGATGAGGAGAGAGAG 6101  
QY 7101 ATTAACCGCTTGAATAGGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7160  
DB 6102 ATTAACCGCTTGTGAG 6161  
QY 7161 CAGCGCTGATGAG 7220  
DB 6162 CAGCGCTGATGAG 6221  
QY 7221 CTGAGCGCGCAGAGCGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7280  
DB 6222 CTGAGCGCGCAGAGCGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6281  
QY 7281 CAGCGCGCATGAG 7340  
DB 6282 AACCGCGCATGAG 6341  
QY 7341 ATGCGAGCGCGATCCCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 7400  
DB 6342 ATGCGATCCCGAG 6401  
QY 7401 CATGATGATGAG 7460  
DB 6402 CATGCGCAATCAGGTAACGAAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6461  
QY 7461 ATTTGCGAAGAGTGAAGTA 7480  
DB 6462 GGTAGAGAGCTTGGAGAGCA 6481

RESULT 4  
112910 LOCUS 6918 bp DNA linear PAT 26-JUL-1995  
DEFINITION Sequence 13 from patent US 5429939.  
ACCESSION 112910  
VERSION 112910.1 GI:910887  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
1 (bases 1 to 6918)  
AUTHORS  
Mitsawa, N., Kobayashi, K., Nakamura, K. and Yamano, S.

TITLE DNA sequences useful for the synthesis of carotenoids  
JOURNAL Patent: US 5429393-A 13 04-JUL-1995;  
FEATURES Location/Qualifiers  
source 1. 6918  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 30.2%; Score 2616.4; DB 6; Length 6918;  
Best Local Similarity 64.7%; Pred. No. 0;  
Matches 4010; Conservative 0; Mismatches 2166; Indels 24; Gaps 7;

QY 1284 CTGAACATTATTTGCTGCGGGGACGACGCGATGCGTTGGGCGGCGATGCGGCC 1343  
DB 303 CTGAACATTATTTGCTGCGGGGACGACGCGATGCGTTGGGCGGCGATGCGGCC 1343  
QY 1344 GGCACGCTGCGACCGGGCAACGATTCCTGCGCTTGGCTGCTGCGACGACGCGAT 1403  
DB 363 GGTGGCGTGGCAACGGGAAACGATTCGCGCCATGTTGCTGTTGCTGACCGCGCGAT 1422  
QY 1404 ATGGGCTGTGACGTGGCGACGACGCGATCTTGATCTTGGCTGCGGTGCAATGGTG 1463  
DB 423 CTGGGTTGCGCTGTGACGACGATGACGCGATTCGATTTGGCTGCGGTGCAATGGTG 1482  
QY 1464 CAGCGTGCCTGACGATTCCTGACGACATTCATTCATGATTAACGCGCGATGCGACGT 1523  
DB 483 CAGCGGCTTGTGCTGATCTTGAACGATATGCGCTGACGACGATGCGAACTGCGCGC 1542  
QY 1524 GGGCGCGCGGCAATCCACTGTGATATGAGGAAACGTTGCGATTCCTGCGACGCGTGC 1583  
DB 543 GGAAGCGCTTACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 602  
QY 1584 CTACTACGCGCGCTTTGAGTGATTCCTGCGCGCGGTGCTGCGACGACGACAA 1643  
DB 603 TTGCTGATGAACCTTTGGCGTAAATGCGATGACGATGCGCTGACGCGCTGCGAA 662  
QY 1644 GCGGAGCCATTCGCGACGCTCTGCGCGGTGGCGCTGACGAGGACGTTGACGGGTG 1703  
DB 663 AATCGGCGGTTTCTGAACGTCAACGCGCATGCGCATGCGAAGATTGTTGACGGGTG 722  
QY 1704 TTCCAGGATCTGATGACGCGACGACGACGAGTCCGGAAGCATCACTTGACCAAT 1763  
DB 723 TTCAAGGATCTGCTGAAGGGGATTAAGCGCGACGCGCTGAAGTATTTGATGACGA 782  
QY 1764 GAACTGAAAACAGCGCTCTGTTTGGCGACGCTGACGATGGCGGATGGCGCGAT 1823  
DB 783 CACTTAAACAGCAGCGCTGTTTGGCTCTCAATGACGATGGCGCTGATTTGTTGCA 842  
QY 1824 GCGTCACTGAGTACGTCAGCGTTTAAGCTATTTTGGCAGAGATTTAGTCAAGCTTTC 1883  
DB 843 GCGTCCAGCGAAGCGCTGATTCGCTGACATCGTTTCACTTGAATCTTGGTCAAGCA 902  
QY 1884 CAGTTATGACGACCTGCGGAGATGGCTTAAGCAGACGCGCAAGACTGTCAAGAT 1943  
DB 903 CAGCTGTGACGATTTGACGATGACGATGACGACGCGTAAAGATTAAGCAATCAGAC 962  
QY 1944 CAGGGAACATCCAGCTGTGATGAGATGCTGAGCGCGGAAGGGGTGAGCTGCTGCG 2003  
DB 963 GCGGTAATTCAGCGGTGATCAATCTGTAGGCGCGGATGGAAGAACGTTGAGA 1022  
QY 2004 GACCATTAAGCAGCGCGATGACACACTTGGCTGCGCTGCGCTGCGGTGTCGAC 2063  
DB 1023 CAGCATCTTCAGCTTGGCAGTGAACATCTCTGCGCGCTGCGCAACGCGGACGCGACT 1082  
QY 2064 CGTCAATATATGACGCGCTGTTTAATCAACAGCTGCGATGTTCACTGAAGCGGTGCA 2123  
DB 1083 CAACATTTTATTCAGCGCTGCTTGAACAAAACGCTGCTGCTGCTGCTGCTGCTGCT 1140  
QY 2124 TACCTATGAGGAGCTTTTGGCTTATGCGCGACGCTTACAGCACTTTACGCAATG 2183  
DB 1141 ---GATGAGCACTTTGCGGCGATGCGACCGCGCTTTTACAGCACTTTGCGCAATTA 1197  
QY 2184 AGGCGCTGCGCAACGCTGCTGCGCGGAGATGCACTTATTCAGCAAGTG 2243

DB 1198 AGAATCTGCTACAGAACTGCTGCGCGGCTGATGCGGTGACCTTTATTCAGCAATAG 1257  
QY 2244 ATGACAGCACTTGTCTGACGACGACGCGATGCTTTGTGGCTGCGCGACGCGAC 2303  
DB 1258 ATATTAACACTGATGATGAGGAAACGATGATGATTTATTCGCTGCGGACAGACGCC 1317  
QY 2304 ATCTGCGGAGATGCTGCTGACGAACTGACGCGGTGCGGACGCGGCGGCTGCG 2363  
DB 1318 ATCTGCGGAGATGCTGCTGACGAACTGACGCGGTGCGGACGCGGCTGCG 1377  
QY 2364 TGTTCGCGTATTCAGATCTGCGACGACGACGCGATGCTATGCTATGCGCGGAC 2423  
DB 1378 TGCTGAAGCTATCAATGAAATGCGCGGACGACGCGATGCTGCTGCGGAC 1437  
QY 2424 CCGTGTGCAACGCGTGCAGTGTGATGCGGTGATGCGGATGCGGATGCGGATGCG 2483  
DB 1438 AGGCAATTAACGATTCGCGCGGTGATGCGGTGATGCGGATGCGGATGCGGATGCG 1497  
QY 2484 GTCTGTGCGCAAGGCTTACAGCTGCGGTGCTGCTGCGGTGCGGCGGCTGCG 2543  
DB 1498 CCGTGTGCGCAAGGCTTACAGCTGCGGTGCTGCTGCGGTGCGGCGGCTGCG 1557  
QY 2544 ATGCGAAGCGGCTTACAGCTGCGGTGCTGCTGCGGTGCGGCGGCTGCG 2603  
DB 1558 ATGCTGAACGATATGCGCGGTGCTGCTGCGGTGCTGCGGTGCGGCGGCTGCG 1617  
QY 2604 CCGTGAAGGCTATACGCGCGACGATGCACTTACAGCGCGATGCGGTGCG 2663  
DB 1618 CTGCGCAAGTATGCGCGCGACGAAATTTATGATGCTGCTGCTGCTGCTGCTGCT 1677  
QY 2664 CTGCTATGCTGCTGATGCGCGGCTTGGCGCTGCGCGGACGCGGCTTACATGAT 2723  
DB 1678 GTGCTATGCTGCGCAACGACGACGACGATGCGGCTTACGCGCGGACGCGGCTTAC 1737  
QY 2724 GTCTGCGCGCTGCGGCAATCAGTCACTGCGGCTTGTGATTTTCCAGCGAC 2783  
DB 1738 GTTTTTCGCACTGCGGCAATCAGTCACTGCGGCTTGTGATTTTCCAGCGAC 1797  
QY 2784 AACTGCGAGCTGCTATCAAGCGGTGCTGCGGCTTGTGATTTTCCAGCGAC 2840  
DB 1798 CGTTACCGGCTGCTTATGCGGCTGCGGCTTGTGATTTTCCAGCGAC 1857  
QY 2841 CGCTGCGCGACCTTGGCGACGCGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCT 2900  
DB 1858 CTTCATTCGCTTATTTTATCTCTGCAAAAACCGCGATTTTGGCTGCTGCGGAC 1917  
QY 2901 TACAGGAGATGCTTGGCTGCTTCTGATCTGCTGCGGCTGCGGATCAGGAC 2960  
DB 1918 TTCAAGGACCGCTTATGCGGCTTGTGATTTTAAACGATGATGAAAGCTTGAAGAA 1977  
QY 2961 TGTGCTGTGTGAGCACTGTGCGGCTGACGCGGCGGACGACATCAGCTGAC 3020  
DB 1978 GTGAGCTGCTGTTAGCGCACTGTGCTGCTTACGAGCTCTGATGAAAGACTGCGCG 2037  
QY 3021 TGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3080  
DB 2038 GAAGCGGATACAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2097  
QY 3081 AACTGTTTATCACTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3140  
DB 2098 AGCTGCGATCACCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2157  
QY 3141 CGATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3200  
DB 2158 CCGTTTATGAGGCTTCCGCTGCTTGTGATCAGCGCGGCTGCTGCTGCTGCTGCTGCT 2217  
QY 3201 ACGCGCTGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3260  
DB 2218 ACGGATGCGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2277  
QY 3261 AACAGTGTGAGTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3320



Dh 2278 GTTCATTGCTGACCAAGCTGACCTTTTCAGCAGCGCATGGCAAAATCCAGACGCCCTTC 2337  
Qy 3321 AGCGGCGCGGTGCTGTAACGGCGCGGTGATATTTGTGACAGCGCTGTGTACGAC 3380  
Db 2338 GTTTGGCAGGGGGGACCAATGGCGCTGCGATATCATTTAGACAGATTATGTC -CACGGT 2396  
Qy 3381 AAATGCTGCTGGCGAGGCGCACTGATGCGCGCACTTATGATGTCATTTCTGTGCGTGC 3440  
Db 2397 CAGCTGTCTTAATGTGGAGGCGGCTATGCAACCGCATTTATGATTCGATTCCTGTGGGGC 2456  
Qy 3441 CGGCTGTGCTAACGGGCTGATGTCGTGCTTTACGCGAGCTGACGCCGCACTTAAAGT 3500  
Db 2457 TGGACTGCGCAATGAGCTTATGCGCTGCTGCTTACAGCAGACCACTGATATGCTAT 2516  
Qy 3501 TTTCCTACTGAGATCAGGCGGCGCGCGCATATATCTGTGCTTCCATCCGCA 3560  
Db 2517 TTTCCTATTCAGACCGCGCACCCGAGCGGGGAGATATACGATGATATTCACAGCA 2576  
Qy 3561 AGACGTGACGAGACGCGAGTTTCGTGCTGACAGCGCGCTTTCCGGCGCGTGGCCCG 3620  
Db 2577 TGAATTGACTGAGAGCCAACTGCTGTGATAGCTCCGCTGCTGATATCCTGCCCCGA 2636  
Qy 3621 TTATCAGATGACGTTCCCGACCTGCTGCTGCGCAGCTGATGTAATTTGCTGATTC 3680  
Db 2637 CTATCAGGTACGCTTTCCCAACCGCGCTGTAAGCTGAACAGCGGCTACTTTTGTATTA 2696  
Qy 3681 CTGCGAGATTTTCCCGGCACTTACAGCAGGTGCTGCTGCTGCGCGCTACCGCA 3740  
Db 2697 TTCTCAGGCTTCGCTGATGTTTACAGCAGATTTGCGCGCACTTGTGAATGATAC 2756  
Qy 3741 GCCGCTCAGCAGGTCTCACCCACCGGGGTGAGCTGCGAGTGGCGGATGTTACAGGC 3800  
Db 2757 CGCGGTGCAAGAGTTAATGCGAATCTGTTCCGTTGAAAAGGCTCAGGTTATCGGTGC 2816  
Qy 3801 GCAGCGGCTGATTTACGAGCGCGGCTGACGCCGACCCGATTCGACGCTCGCTATCA 3860  
Db 2817 CCGGCGGTGATTAACGCGCGGGGTATGCGGCAAAATTCACACCTGAGCGGCTTCCA 2876  
Qy 3861 GGCAATTTGTGGTGAAGAGTGGCAACTGGCGCGCGATGGCGCTGACAGCAAGCAATTT 3920  
Db 2877 GGCGTTATTTGGCAGAGATGGCAATTTGAGCACCCGATGTTATGCTCTCCATTTAT 2936  
Qy 3921 GATGAGCGCGACGCTGATCAGACAGAGGTTATCGCTTTGTTTACACCTTGGCGCTCAG 3980  
Db 2937 CATGATGCGCACGCTGATCAGAAAATGTTATCGCTTGTGTAACAGCGCTGCGCTC 2996  
Qy 3981 TGGCAGCGCTTACTGATGAAGATACCCTATCATCAACATGCCAGCTGATGCCGC 4040  
Db 2997 GCCGACCAATTTGTAATTTGAAGACGCACTATATGATTAATGCCATTAAGATCCTGA 3056  
Qy 4041 ACAGGCGCGCGTCACTTACGATTTATGCCACAGCGCGGCTGGAATTTGGCGCAGCT 4100  
Db 3057 ATGCGCGGCGAAATTTTGGACTATGCGCGCAGACAGGATTTGGAGCTTCAGACACT 3116  
Qy 4101 GCTGCGGAGAGACGAGCTGCTGCGCATCAGCTGACGCGGANTATGATTCAGTTCTG 4160  
Db 3117 GCTGCGAAGAAACAGGGCGCTTATCCCATTACTCTGCGGCAATGCCAGCGATCTG 3176  
Qy 4161 GCACACAGCAGACGGGCACTGTCAGAGGGCTGCGCGCGGATGTTTACAGCCACAC 4220  
Db 3177 G--CAACAGCGCCCTGCGCTGTAGTGAATTAAGTCCGCTGTTTCCATCTTACAC 3233  
Qy 4221 CGGTACTCGCTGCGCGCGGTGCGGTGGCGGAGAAATTTGCCAGCAGCGCTGCCGC 4280  
Db 3234 CGGCTATTTCACTGCGCGCTGGCGGTTGCCGTGGCGACCGCTGATGTCATTGATCTT 3293  
Qy 4281 CGAGCTCACAACGCTGAGCACTGTCATGAATCTTTGCCGTCAGACACTTGGCGGCA 4340  
Db 3294 TACGTGGGCTCAATTCACCATTCATTCAGCAATTTTCCCGGAGGCGGTGGCAGCAGCA 3353  
Qy 4341 GCGCTTTTTCGCTGTGTTAAATGCAATGCTGTTTCTTCCGAGAGGCTTGAACGCGCTG 4400  
Db 3354 GGGCTTTTTCGCAATGCTGATGCAATGCTGTTTTCGCGGAGCCGCGCAATTCACGCTG 3413

Qy 4401 GCGGTAAATGCAAGCTTTTACCGGCTTACGCGGATGATTAAGCGGCTTTTACGCGG 4460  
Db 3414 GCGGTATACAGGCTTTTATGATTTACTGTAAGATTTATGTCCTGTTTATGGGG 3473  
Qy 4461 GCACTGCGGCTCAGCAGATTAAGCAGCATTTCTGTGCGGCAACCGCGGCTCTCGG 4520  
Db 3474 AAAATCAGCTGACCGATGGGCTAGGTATTTCTGAGGGGCAACCGCTTCCGATTT 3533  
Qy 4521 CGAAGCGCTGCGGCTTATGATGATCTCTCCGTTACAGGGAGAAATATGAAGC 4580  
Db 3534 AGCAGCATTCAGGCACTTATGACATCACTCATCTTAAAGGCACTA---CATGAACCA 3590  
Qy 4581 ACCATGTAATGGCGAGGCTTCCGCTGCGCTGCGCATTTGCTTGAAGCGGC 4640  
Db 3591 ACTACGTAATGCTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3650  
Qy 4641 GCGCTGCGGTCAAGCTGCTGGAACAGCGCATTAAGCTTGGCGGCGGCTATGCTAT 4700  
Db 3651 GGAATCCCGCTTACTGCTTGAACAAGTATTAACCGCGGCTGCGCTTATGCTAC 3710  
Qy 4701 CAGATCAGGCTTTTACCTTTGATGCGGCTCCGAGCTGATTAACGATCCAGGCTATC 4760  
Db 3711 GAGATCAGGGGTTTACCTTTGATGAGCGCCGATTAACGATCCAGTCCAGTGCATT 3770  
Qy 4761 GAGGCGCTTTTACGCTGCGGCAAGCACTGATTAATGTCGACTGATGCGCGTG 4820  
Db 3771 GAAGAACTTTTTCAGCTGCGGCAAGAAACATTTAAAGATGTCGACTGCTGCGGTT 3830  
Qy 4821 AGCGCAATTTTATCGCTGCTGGAAGAGCAGGAGCAGCTGACATCAACAAATCAG 4880  
Db 3831 AGCGGCTTTTACCGCTGCTGCTGGAAGTCAAGGAAGCTTTAATTAACATCAATCA 3890  
Qy 4881 GCGGAGTGAAGACAGATTTGCCACTTTAATCCCAAGATGTCGCGGTTACCGCAG 4940  
Db 3891 ACCGCTTCAGACGAGATTCAGCAATTTAATCCCGCATGTCAGAGGTTATCGCAG 3950  
Qy 4941 TTTCGCGCTTATTCAGAGATGATTTTCTGAGAGGCTATCTGAACCTGGGCACTGACT 5000  
Db 3951 TTTCGAGCTTATTCAGCGCGGCTGTTTAAAGAGCTATTAAGCTCGCTACTGCTCCT 4010  
Qy 5001 TTTCTGATTTCCGCAATGCTGCTGCGGCGCAACGCTGAGTGGCTGACGCGCTG 5060  
Db 4011 TTTTATTCGTTCAAGACATGCTTCCGCGGCACTCACTGCGAAATGCGAGCATGG 4070  
Qy 5061 CGCAGTCTACAGCATGCTGGGAAATTTATCATGACATATCTGCGCGGCTTT 5120  
Db 4071 AGAAGGCTTTACAGTAAGTTGCGAGTTACATGAAAGATGAACATCTGCGCGAGGCTT 4130  
Qy 5121 TCCCTTCACTGTTGCTGCTGCGCGGTAATCCTTTGCAACGTCCTGATCTATACCTTA 5180  
Db 4131 TCTTTCACTGCTGCTGTTGGGCGGCAATCCTTGGCACCTCATTCATTTATAGCTTG 4190  
Qy 5181 ATTACGCACTGAGCGAGCGAATGGGCGTGTGTTTCCGCGCGGCTACCGGTGCGCTG 5240  
Db 4191 ATACAGCGCTGAGCGTGAAGTGGGCGTGTGTTTCCGCGTGGCGCACCGCGCATTA 4250  
Qy 5241 GTTGATGCAATGCGCGGCTTTTCCGATTTTGGGCGGTGAATGCTGCTCAACGCGGA 5300  
Db 4251 GTTCAGGGAGTGAATTAAGCTGTTTACAGATCTGGGCGGAATCGTGTAAACGCGAGA 4310  
Qy 5301 GTGACGAGCTGAGAACCGAGGTTAACCGCATCAGCGGTGTCAGCTGAAAGTGGGCGC 5360  
Db 4311 GTCAAGCATTTGAAAGACAGAGAAACAGATTTGAAGCCGTGATTTAAGAGCGGTGCG 4370  
Qy 5361 CGTTTTCGCGCGCGCGCTGCTGCTGCAATGCTGACGCTGCTGATCACTGATGCTG 5420  
Db 4371 AGGTTCTGACGCAACCGCTGCTGCTGCAATGCTGATGCTGATCACTGATGCTG 4430  
Qy 5421 TTAAGCAGATCTGCGCGCTTAAACGCGGCAACGCTGAAGGCAAGCGGATGAGC 5480  
Db 4431 TTAAGCAGACCCCTGCGGCTTAAAGCAGTCCAAACAACTGAGACTGAAGCGCATGAGT 4490



QY	5481	AACCTGCGCTTTGTAAGCTATATTTGGTCTTATATCATGCGCAACCGAGCTGGCGACAAC	5540
Db	4491	AACCTCTGTTTGTCCTTATTTGGTTTGAATCACCATCATGATCAGCTCGCGCATAC	4550
QY	5541	ACGGTGTGCTTTGGTCGCGCGCTATCGTAATGTATCGATGAGATCTTCAATAGCAGCCAG	5600
Db	4551	ACGGTTTGTTCGGCCCGCGTATCCGAGGCTGATGACGAATTTTAAATCATGATGCG	4610
QY	5601	CTGGCGGAAGATTTCTCGCTGTATCTGATATGCGCCCTGCTCCAGCGATCCGTGCTGGCA	5660
Db	4611	CTCGCAGAGGACTTCTCACTTTATCTGCACGCGCCCTGTGTCAGGATTCCTCACTGGCG	4670
QY	5661	CCGGCGGGCTGGGCGAGTTTATAGTGTCGGGGCGGGCGGCAATCCGGAATCCGCGCA	5720
Db	4671	CTGGAAGTTGGGCGAGTTACTATGTGTGTGGGCGGGCGGCATTATAGGCAACCGCAAC	4730
QY	5721	ATTGACTGGCAACAGAAAGGCGCGCTGCGCATCGCATTTTGTCTTATCTGGAGAG	5780
Db	4731	CTGCACTTGGACGGTTGAGAGGGGCCAAACTAGCGCACCGTATTTTGTGCTACCTTGAGAG	4790
QY	5781	CATATATATGCGGGGTCTGCGACAGCAGTTAAGTACACACCGTATGTTTACGCCGTTGAT	5840
Db	4791	CATTACATGCTGGGTTACGGAGTACGTGTGTACGACACCGAATGTTTACCGCCTTGAT	4850
QY	5841	TTTGGCGGACAGCTGACAGCGGCATCAGGGGCTCAGGTTTGGCTCGAACCCATTTTGAC	5900
Db	4851	TTTGGCGACAGCTTAAATGCTATCATGAGGCTCAGGCTTTTGTGAGAACCGGTTCTTAC	4910
QY	5901	CAAAAGCGCTGTGTTCCGGCCGCATPACCGCGATGCGCATTAATACTTTATCTGGTG	5960
Db	4911	CAGAGCGCTGTGTTCCGCCGCCATPACCGCATTAACCATTTACTAATCTTACTGCTC	4970
QY	5961	GGGGCTGGCAGCATCCCGGTGCGGGTGTGCAAGCGGTGATCGCTCCGCGAAACGACC	6020
Db	4971	GGCCGAGGCAAGCATCCCGCGCAGAGCATTCCTGGCGTATCGGCTGGCAAAAGCAGCA	5030
QY	6021	GCCGACGTGATGTGTGAGAGATCTGACCGGATGAACCAACCGCCGCTGATTGACAGGTCA	6080
Db	5031	GCAGGTTGATGTGTGAGAGATCTGATTTGAATTAATCCGTGTTACTCATATATGCGGTC	5089
QY	6081	CGCAACCATGAGCGCAGGGGCTCCAAAGATTTCCGACAGGCTACCGGCTATTTGATCCTT	6140
Db	5090	--GAAACGATGAGCAGTTGGCTCGAAAGTTTGGCACAGCCTCAAAAGTTATTTGATGCA	5147
QY	6141	CAACGCGCGCAGTACGCTGATGCTGTACGCGCTGTGTGCTACATGTGACGATGTGATG	6200
Db	5148	AAACCGCGCGAGCGTACTGATGTCTACGCGCTGGTCCGCAATGTGACATGTATGT	5207
QY	6201	ATGTGTGACGCTGTGGGCGAAGCGGCACGACGACGCGGTGGCGGATGCAACGGCGGCA	6260
Db	5208	ACGATTCAGAACCTGTGGCTTTTCAAGGCCCGGACGCTGCTTCAAAACGCCAACAACGTC	5267
QY	6261	TGGCGACCTGTCAAAATGGAACCGCGCGGCTTACAGGGGTGCCACATGGATGAACGAG	6320
Db	5268	TGATGCACTTGAGATGAAGAACCGGCACAGGCGCTTACAGGATCGAGATGACAAACCGG	5327
QY	6321	CGTTTGTGCTTTCAGGAAGTGGCGCTGACGCAATCAGCTTCCCGACGAGCTGGCTTTTG	6380
Db	5328	CGTTTGTGCGCTTTTCAAGGAAGTGGCTAATGGCTAATGATATGCCGCCGCTTACGCCGTTTG	5387
QY	6381	ATCATCTTGAAGGTTTGCATGTGATGCGGTGAAGAACTTATATGCTGTTTCGGAGACA	6440
Db	5388	ATCATCTTGAAGGCTTCCGCATGTGATGAACCGAAGCGCATTAACGCAACTGTGATGATA	5447
QY	6441	CGTGCCTTATCTGCTATACGTTGCGCGGTGTGCGGTTTATATAGCGCGCGCTGATG	6500
Db	5448	CGTGCCTTATGCTATACGTTGAGGCGTGTGCGCTTATATATGCGGCAAAATCATG	5507
QY	6501	GCGTACGTGATGAGCGCTACTCTCATACGCGCTGTGATTTGGGCTGTGCGCTTACGCTTAA	6560
Db	5508	GCGTGCGGGATTAACGCAACGCTGGAACCGCGCTGTGATCTTGGGCTGTGCAATTTCA	5567
QY	6561	CGAATATTCACGGGATATCGTTGAGACGCGGAGAATGCGCGTGTCTATCTGCGCAACA	6620

Db	5568	CCAAATATTGCTCGGAGATATTGTGGACGATGGCATGGCGGGCCGCTGTATTCTGCGGGCAA	5527
Qy	6621	GCTGGCTTGAAATGAGGCGCGAAGCTGAGCGCCGCGCCAGCTTGCCGATCCGCAACATGTGCGCAG	6680
Db	5628	GCTGGCTGGAGCATGAAAGGCTCTGAAACAAAGAAATTTATGGCGGCACCTGAAAAACCGTCAAG	5687
Qy	6681	CGCTGGCCCGGCTGGGACGCGGCTGGGTGGCGGAGCGCGAGCCGTACTATCAATGACAGCGC	6740
Db	5688	CGCTGAGCCGATATGCGCCCGCTGTTTGTGTGACGAGAGAGAACTTACTATTTGTCTGCA	5747
Qy	6741	GCAAGCGGCTGCCAGGATTTGCCGCTCCGTTCCGGCTGGGCGATGCCACCGCGCGCGAGC	6800
Db	5748	CAGCCGCGCTGGGAGGGTTGCCCTCGCTTCCGCTGGCAATGCTAACCGCGAAGCAGG	5807
Qy	6801	TTTATCCGGGAAATTTGGGCTTAAAAAGTGCAGCATGCGCGTGCCTCGGCGCATTGGGATACGGGCC	6860
Db	5808	TTTATCCGGAATAATGAGTGTCAAAAGTTAACAAGCCGGGTCAAGCAAGCTTGGGATACAGCGGC	5867
Qy	6861	AGCGCACACAGTAAAGGAGGAAAGGTGGCGCTGCTGGTGAAGGAGGCGCGCGCTGCGGCTTAA	6920
Db	5868	AGTCAACGACACCGCCCGGAAAAATTAACGCTGCTGCTGGCGCGCTCTGTGTCAGGCCCTTAA	5927
Qy	6921	CTTTCGCGCTTGCTCATCCGAGGCGGCTCTGCGGATCTGTGGACGCTTCCGCTTGAAC	6980
Db	5928	CTTTCGCGGATGGCGGCTCATCTCTCCCGCCTTGCGCATCTTGGCAGCGCCG-----C	5981
Qy	6981	ACGACGCGCCATTTGGCGCTGGCGGACGCGCTGCGCTGAGCTTGTGCAACGGTGGCGGCTAAAG	7040
Db	5982	TCTAGCGCCAGTGTTCCTCCGAGCGTGGCTGAAGTTTGTGACAGGGGCGCGCATAGAG	6041
Qy	7041	AAAGCCAAAGGAAACGAGGCTTTCGCCGTCCCGCACCGCATGATGATGCGGTGGCGCAT	7100
Db	6042	GAAAGCAAAAGAAACACACACTTCTTTGCCCTGACGCGCGATGATACGATCGGTGCGCAT	6101
Qy	7101	ATTACACCGCTTACAGATAGCCTTTTGGCTGGAAATTAAGCGAAACGGCCAGGCTTATATGAC	7160
Db	6102	ATTACACCGTTTGGAGTGGCCCTTGGCGGTGAAATTAACCGAATGGCCAGGTTGATATCAC	6161
Qy	7161	CAGGCGCTGATNGCACCATTAATAATGAGAGGCGCATAGTGGTCTATGCGACGCGCAATTCGA	7220
Db	6162	CAGCGCGTCTGTGACCATTAATAATGAGTAAATCCATAGCCGTCTATCTGTGCGCAATTCGA	6221
Qy	7221	CTGCAAGCGGCGCAGAGCGCGTTGACACCCAGCCCAATAGGCGCAATCGACAAACAGCGCAA	7280
Db	6222	CTGGAAGCGGCGCACTTCTGTACTGCCAGATTAATCAGCAGAGATCGATATATGACAGCAA	6281
Qy	7281	CACCAAGGCGATCAGGCTGCTTGAAGTCTCTCATGTGTGTTCTCATGATGTCGACAA	7340
Db	6282	AAACCAAGGCGATTAAGATCGTTAACTTCAAGCAGCACTTTTACGGCGTTCAATGATGTGAAG	6341
Qy	7341	ATGCAAGCCCATCTCCAGCGCATATGATATTTATTTGAGACAGCGCCGCGAGCATCTC	7400
Db	6342	ATGCAATCCCAACCCAGCGGTGACATGATATTTGTGTGCGCAGTGCAGCAATCATCTTC	6401
Qy	7401	CATCAATATCAAGGTAGCCAGCAGGATTAAGCGCATTCATTAACCAAGCATCATTTGCTCC	7460
Db	6402	CATGCGCAATCAAGGTAAAGAAACGATCAGGGCATTCCAATTCACAAACATTAATTTCTCC	6461
Qy	7461	ATTGGCAAGATGAGAGTA 7480	
Db	6462	GGTAGAGACGTTGTGGCAGCA 6481	

RESULT 5  
ERMCR1A 6965 bp DNA linear BCT 11-APR-2001  
LOCUS  
DEFINITION  
ERMCR1A  
Pentose agglomerans (crx) CrtX (crxX), glycane cyclase  
(crtY), phycone dehydrogenase (crtI), and prephytoene  
pyrophosphate synthase (crtB) genes, complete cds.  
ACCESSION  
M50638  
VERSION  
M50638.1 GI:148393  
KEYWORDS

SOURCE ) Pantoea agglomerans  
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
REFERENCE 1 (bases 1101 to 1300; 4301 to 4410; 4641 to 4750; 4881 to 5000;  
5031 to 5140; 5561 to 5660; 6031 to 6160)  
AUTHORS Liu, S.T.  
TITLE Carotenoid-biosynthesis genes as a genetic marker for the purpose  
of gene cloning  
JOURNAL Biochem. Biophys. Res. Commun. 195 (1), 259-263 (1993)  
MEDLINE 8395826  
PUBMED 93371414  
REFERENCE 2 (bases 1 to 6965)  
AUTHORS To, K.Y., Lai, E.M., Lee, L.Y., Lin, T.P., Hung, C.H., Chen, C.L.,  
Chang, Y.S., and Liu, S.T.  
TITLE Analysis of the gene cluster encoding carotenoid biosynthesis in  
Erwinia herbicola Bho13  
JOURNAL Microbiology 140 (Pt 2), 331-339 (1994)  
MEDLINE 94236237  
PUBMED 8180698  
FEATURES  
source  
1. .6965  
/organism="Pantoea agglomerans"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:549"  
776..6687  
/product="CrtE"  
926..1834  
/gene="crtE"  
926..1834  
/gene="crtE"  
/codon\_start=1  
/transl\_table=11  
/product="CrtE"  
/protein\_id="AAA21260.1"  
/db\_xref="GI:148394"  
/translation="MTVCAGKGVHPTSSAAEQLLADIDRLLEQLPYEGERDFYGAAM  
REGALAPGKRIRPMILITLARDIGCAVSEGLIDACAVEVHAASLIIDMPCMDA  
OIVGGRPTVHCQYGHVAILLAVALLSKAPFVINDAGCTPLAKRAVSELNAGMO  
GLVGGQFQDLSEGRSAAEILMTNFKITSLTLCASQMASTYVANNSSERRDYLHP  
SLDGOAFQDLDDTDGMADYKQDINDEKSTLVNLLGSRAVERLRHSLHLSEHL  
SAACNGHSTQFIQAMFQKLAAS"  
1844..3139  
/gene="crtX"  
1844..3139  
/gene="crtX"  
/codon\_start=1  
/transl\_table=11  
/product="CrtX"  
/protein\_id="AAA21261.1"  
/db\_xref="GI:148395"  
/translation="MSHPAATAPPYSYHVALLENLAQELVARGHQVTFIQODTKPLI  
QPHIGFAVGEESHPGSLRLHLAHLAPSPNLRLINMARTSDMLCIALPRATN  
DLADIGVLDMEPAGALVABALGIPVSVACALPLNRPDLPAVWFDGTDAR  
ERYASEKIVMLMRHDLVIAHSHHSEKAPRDLKCFGLPLAQISQILPELDPRK  
ALPCFHAVPLRETQSTSTSPYFSESEKPRIPASIGTQGHRYGFKAIYACEE  
IDGQILLAHCGRIIRFOABELABSRHNOVDADSOALITHGGMNTLDLI  
NRTPLALPLAFDQGVASKRIVSHGIRKRSRTTSALARRHRLITNDYQKRT  
KIQTALRLAGGMAADIVEQAMRTGQVLGSGVATL"  
3123..4271  
/gene="crtY"  
3123..4271  
/gene="crtY"  
/codon\_start=1  
/transl\_table=11  
/product="dycopene cyclase"  
/protein\_id="AAA21262.1"  
/db\_xref="GI:148396"  
/translation="NRPAYDILVAGANGLIARLDOQDPDKRILLIEAARQGN  
HTWSPHADLTVEQHRWAPLVVYHWPVYVRFPTRRKLNISGFSVTSQFAELQO  
KFGQHLWISRAVAEVAHADVRLNNGQVISAQVINDGRKTNLSALNVGFQIOWMR  
LSKPHGLSSPIMDATVDQNGYRFVSLPISATELLIEDTHYIDNATLBERARONI  
RDYAAQDWMQGLTLREEGALPITLTGDSATFWQOQPLACSGLAGLHFHTGSLP

gene  
CDS  
LVALADRLSALDVFSSSIHOAITFPAHRNQOQFFRMNLRLFLAGPADSRWRVM  
ORFYGLPEDLIRSFYAGKLTLDRLRLISGKRPVPVLAALQAIMTTR"  
4283..5761  
/gene="crtI"  
4283..5761  
/gene="crtI"  
/codon\_start=1  
/transl\_table=11  
/product="phytoene dehydrogenase"  
/protein\_id="AAA21263.1"  
/db\_xref="GI:148397"  
/translation="MKPTTVIAGAGGGLALAIRLOQAIPVILLIEORDKPGRAYYE  
DQGFPTDAPPTVITDPSAIEELFTLAGKQKDYVELPVAPEYRLCWSGKXFNVDN  
QAOLEAOIOQFNPDVEYGRQFLDYRAVFKEGYLGTVPFLISFRDMLRAAQLAKL  
QAMRTYVSAVSIENDEHLROAFSPHSLVGNPFAVSSITLTLHLEREMGWVPRG  
GTICALVKGMITKIFQDQGGVAVLNAAKYSHMETTGTEAVHLDGRFPTRAVANSADV  
VHTYRLDSQHPAAVQSKKLQTKRSSLNLPVLYEELNHNHQLAHNTYCFGRYREL  
LHEIFNHDGLADFLDYLAHPCTVDSLSLEEGGSYYVLAIPVHLGTALDWTGSPR  
LRDRIPFYLEQHMYPGLRSOLVYORMTFPTFPDQLANAQSAPSVPEVLTOSAWERP  
HNDRKTINMLYLVGAGTHPGAGIPVIGSAKATAGIMLEDLI"  
5797..6687  
/gene="crtB"  
5797..6687  
/gene="crtB"  
/codon\_start=1  
/transl\_table=11  
/product="prephytoene pyrophosphate synthase"  
/protein\_id="AAA21264.1"  
/db\_xref="GI:148398"  
/translation="MAVGSKSPATSKLPDAKTRRSVLMYKMKRCHDDVIDQDTGF  
QTDQPSIQTPBQRLQMLEMKTROVAVGSOQWHEPFAFAFEVANAHDIAVAVADHLEG  
FAMDVREAEYIQLDPTLRYCYHAGVGVGMAQIMQVNDATLDRACDGLAVFLNI  
ARDIVDAHVGRCYLPASMLENGLNKENYAAENQALSRILRLVQSAEPYLLCAT  
AGLAGLPLRSAMVATAKOVYRKIGVKGVAQGEAMVDHNGSTTPPKLSILLMASSQA  
ITSRMHPHPRPHLMORP"  
ORIGIN  
Query Match 28.7%; Score 2486.2; DB 1; Length 6965;  
Best Local Similarity 64.4%; Pred. No. 0;  
Matches 3852; Conservative 0; Mismatches 2103; Indels 24; Gaps 8;  
QY 1284 CTTGAACATTATTTGCTGCGCGGAGACAGCGCGATCGCTTGGCGCGGCATGCGTCC 1343  
DB 1004 CTTGAGCAGTTATGCTGTCGAGGAGAAACGGGATTTTGTAGCTGCGCGGAGTCCGAG 1063  
QY 1344 GGCACGCTGGCACCGGCAACGATTCGTCGCTGTCGTCGCGGAGCAGCGCAT 1403  
DB 1064 GGCAGCTGGCACCGGATTAACGATTCGCGCATGCTGTTGCTCACAGCAGCGCAT 1123  
QY 1404 ATGGGCTGTACGTCGCGCAGACAGCGGATCTTGATCTTTCGCTGTCGATCGAAATGTTG 1463  
DB 1124 CTGGGCTGCGCTGTCACCGCAGGAGGATTCGCGATTCGCTGCGCGCGGCGGAAATGTTG 1183  
QY 1464 CACGCTGCTCATGATCTGTCGACGACATTCATCAATGATTAACGCGCGGATCGACAT 1523  
DB 1184 CACGCGGATCATGTTGATCTTGATGATATCCCTGTATGATATGAGGACGATCTGCGC 1243  
QY 1524 GGGCGCGCGGATCACTGATGATATGAGGAAACGTCGCGATCTGCGAGCGGTCGCG 1583  
DB 1244 GGGCGCGCTACCGTTACCTGACGATACGCGAGATGCGCAATATCTGGCGGCGGTCGCC 1303  
QY 1584 CTACTCAGCGCGGCTTTGAGGATGATTCCTCGCGCGCGGATCTGCGAGCAACGACAAA 1643  
DB 1304 TTACTGAGCAAAAGCTTTGGGGTCAATTCGATGACAGATGCGCGACGCGGCTGCGAAA 1363  
QY 1644 GCGGACCATTTGCGGAGCTCTCTCGCGTGGGCTGCGAGGAGCTGTTACAGGCTCAG 1703  
DB 1364 AACCGGCAAGTGCAGATTAATCAACGCGCATGTTGCAAGGTTTGTACAGGGCTCAG 1423  
QY 1704 TTCCAGGATCTGATGACGCGCAGACAGCGCGAGTCGGAAGGCATCACCTGACCAAT 1763  
DB 1424 TTTAAGGATCTGTCAGGAGGAGATTAACGCGCGCGCTGACGCAATTTTGTATGACGAC 1483

1764 GAACGAAAAACGAGCTGCTTTTCGCGCACGCTGACGATGAGCGGCGAT 1823  
1484 CACTTTAAACCGACACCTGTTTTCGCGCTATGAGATGAGCTTAATTTGCGCAAT 1543  
1824 GCGTCAGTGAAGGATGCTGACGCTTTTAACTATTTTTCGCGAGATTTAGGTCAGCTTTC 1883  
1544 GCGTCAGGAGAAAGAGCGGATTAACCTAACACGTTTTTCGCTTATCTCGGCGCAAGCTTT 1603  
1884 CAGTTATCTGAGCACTGCGGATGCTTAAGACACCGCGCAAGACTGTCAAGAT 1943  
1604 CAACGTGTGACCACTTAACGATGAGCTGACACGCGTAAAGCACTCAATCAAGAT 1663  
1944 CAGGCGAAATTCACGCGTGTGACGATGCTGAGCGCGGAGAGGCGTGAAGCTGCTGCGC 2003  
1664 GAAGGCAAAATGACCTGCTGCAATTTTATTAAGCTTCAGAGCACTTGAAGACGCTGCGG 1723  
2004 GACCATTTAAGACGCGCGATGACACACTTTCGCGCTGCGCTGCGCTGCTGCGCGCACC 2063  
1724 CACAGTGTGATCTCGCGAGTGAAGCACTCTCAAGCGCGGTGTCAAAACGCTCAATCTACT 1783  
2064 GGTCAATATATGACGCGCTGTTTATCAACAGCTGCGATGTTCAACTGAAGCGGCTCA 2123  
1784 CAACAATTTATCAGGCGCTGTTTGAACAAAACCTCGCTGCGCTGAAGATGCT-- 1841  
2124 TACCTAATGGGCAATTTTCCGTTATTTGCGCACCGCTCTACAGCACTTTCACGCAATTC 2183  
1842 ---GCAATGAGCACTTTCGCGCGATGCTCGCTTTTTCAGCAAGTCTGTCACCTG 1898  
2184 AGCGCTGCGGCAAAACGCTGCTGCGCGGAGCACTGCGCATCACTTTATTCAGCAAGT 2243  
1899 AAAACCTCGCTCAGAGATTAAGCTGCTGCGGTATCAGGTGACCTTTATTCAGCAAG 1958  
2244 ATGACGACACTTGTGAGGAGCGAGCGCAATGCTTTGTCGCGCTGCGGAGCGACG 2303  
1959 ACACTAAACCACTCATTCAGCATCCCAATCGGATTTCAATGCTGAGGAGAAAGAGCC 2018  
2304 ATCTGCGGATGCTCTCCAGGCAATCAGCGGCTGCGCGACCGGCGGCTGCTGCG 2363  
2019 ATCCGCTGCTCATTTATGCGCTGCTGCACTTAAGCGGCCATCTCTGCGGCGCTCA 2078  
2364 TGTTCGCGGATTTACAGATCTGCGCGACGACGATATGCTATGCGCGCAATGCGCG 2423  
2079 TGCTTGCGCTCATTAATGAATGAGCGCGCACGCGATATGCTGCTGCGCTGCCA 2138  
2139 GAGGCTTTAAGATTGCGCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2198  
2424 CCGTGTGCAACGCGCTGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2483  
2484 GTCGTGTGAGAGAGGCTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2543  
2199 CTCTGTGCTGAGAGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2258  
2544 ATGCGGAGGCGGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2603  
2259 ATGCGGAGCGGATTTCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2318  
2604 CCGTGTGCAACGCGCTATCAGGCGAGCGAGTGAATCTAAGCAACGCGATGCTGCTGCTG 2663  
2319 CAGGTAGCGGCTATGAGCGAGCGAATAATTTATGCTGCTGCTGCTGCTGCTGCTGCTG 2378  
2664 CTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2723  
2379 GCGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2438  
2724 GTCGTGCGCGCTGCGCGCAATCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2783  
2439 GTTTTTCGCTATGCGCAAAATCAGTCACTGCTTATTTCAAGAACTGCTGCTGCTGCTGCT 2498  
2784 AACTGCGAGCTGCTATCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2841  
2499 CGGTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2558  
2842 GCTGCGCGACCTGCTGCGAGCGCTGCG--CAGCGCGGTGCTGCTGCTGCTGCTGCTGCTG 2900

2559 CTTGCGCGCTTATTTTTCGCGCTGAGAAAGCGCGAATTTTCGCTGCTGCGGACG 2618  
2901 TACAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2960  
2619 TCGAGGCTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2678  
2961 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3020  
2679 GTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2738  
3021 TCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3080  
2739 GAAGCGCTACAGCGAGCTTGTGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2798  
3081 AACTGTTTACTCACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3140  
2799 AGCTTGCATTTACCTGAGGAGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2858  
3141 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3200  
2859 CCGTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2918  
3201 AGCGCGTGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3260  
2919 ATGAGATGAGAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2978  
3261 AACGTTGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3320  
2979 GTGCTATGCTGACCAAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3038  
3321 AGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3380  
3039 GTTTAGCGCGCGGCAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3098  
3381 AAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3440  
3099 AGCCGCTTTTGAAGGAGGAGCGCT--ATGCAACCGCATTAATGCTGCTGCTGCTGCTGCT 3157  
3441 CCGCT 3500  
3158 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3217  
3501 TTTGCTATGAGAGTCAAGCG 3560  
3218 TGTGCTTATGAGAGCG 3277  
3561 AGAGTCAAGAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3620  
3278 TGAATTAACCGAAGCGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3337  
3621 TTTATGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3680  
3338 CTATGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3397  
3681 CTGAGAGATTTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3740  
3398 TTTGCGAGCGCTTTCGCGAGGCTTTTACACAGAAATTTTGTGCTGCTGCTGCTGCTGCTG 3457  
3741 GCGGCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3800  
3458 CGGCTGCGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3517  
3801 GCAGGCGGCTGATTTGAGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3860  
3518 CAGCGCGGCTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3577  
3861 GCGATTTTGTGCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3920  
3578 GCGGCTTATGAGCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3637  
3921 GATGAGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3980

Db 3538 TATGATGCCACGGTCATGACGAAAAATGTTATGCTTTGTTTACAGCTCCGCTATC 3597  
 Qy 3981 TGCAGCGTTTACTGATTTGAAGATACCCACTATCAATCAACCATGCCAGCTGAATGCC 4040  
 Db 3698 TGCACAGAACTGTTAATCGAAGATACCCACTATATGATTAACGACACTGAACTGGA 3757  
 Qy 4441 ACAGCGCGCGGTACATTAAGATTAAGCCACAGCGCGGTGGAATTTGGCCGAGCT 4100  
 Db 3758 ACAGCGCGGACAAAATATTCCGGAATTATGCTGCCAGAGGATTTGCGCTTCAGACTCT 3817  
 Qy 4101 GCTCGGAGAGAGACAGCGCTGCTGCGGATCAAGCTCAGCGGATATCATAGTTCTG 4160  
 Db 3818 GCTTCGTAGGAGAACAGCGCGGATTTGCGAATACCTTAACGCGGATATGACACCGGTTTG 3877  
 Qy 4161 GCACACAGACAGCGGACAGCTGAGCGGCTGCGCGCGGACTGTTTACGCGACAC 4220  
 Db 3878 GCACACACAGC---CTTATGCTGACAGTGCAGCTGCTGCGGACTGTTTACACCCACAC 3934  
 Qy 4221 CGGTTACTGCTGCGCGCGGCTGCGGCTGCGGAGAAATGCGACAGACGCTGCCGCG 4280  
 Db 3935 CGGTTACTGCTGCGCGCGGCTGCGGCTGCGGAGATGCGCTGAGTGCCTGATGTTCTT 3994  
 Qy 4281 CGAGCTCACACAGCTGAGCCACTGATCGAATCCTTTGCCGCTGACGACTGCGCGAGCA 4340  
 Db 3995 TACATCGTCTCAATTCATGAGGCGATTACCCACTTTGCCACGAGGCGCTGCGAGCA 4054  
 Qy 4341 GCGCTTTTCCGCTGTTAAATCGATGCTGTTCTTCCGCGAGCGCTGAACGCGCTG 4400  
 Db 4055 ACAGCTTTTCCGATGTTGAATCGCATGCTGTTTTCGCGGACCGCGCGATTCACGCTG 4114  
 Qy 4401 GCGGCTATGACGAGGTTTTCACCGGCTGACGCGGATGATTAAGCGCTTTTACGCGCG 4460  
 Db 4115 GCGGTTATGACGCGCTTCTATGCTTACCTGAAGATTAAATCTCCGTTTATGCGGG 4174  
 Qy 4461 GCACCTGCGCTCAACGATTAAGACAGCATTTCTGCGGCAACCGCGGCTCTCGG 4520  
 Db 4175 CAATCTACAGCTGACCGATCGGCTACGTAATCTGAGCGGCAACCGCTGTTCCGGTAT 4234  
 Qy 4521 CGAAGCGCTGCGGCTTGAATGATGACCTCTCCGTTACAGGGAAGAAATTAATGAACG 4580  
 Db 4235 AGCGGCAATTGACGCAATTAAGACATCATGTTAAAG---GCAACACATMAACA 4291  
 Qy 4581 ACCTATGATTTGCGGAGGCTTCCGTTGCGCTGCGCTGCGGATTTGCTGCAAGCGG 4640  
 Db 4292 ACTACGGAATTTGATGAGGCTTTGCTGCTGCGCTGCGCTTGAATTCGCTGCAAGCGG 4351  
 Qy 4641 GCGGCTGCGCTGACGCTGCTGGAACAGCGGATTAAGCTGCGGCGCGCTATGCTAT 4700  
 Db 4352 GGGATCCCTGCTTACTGCTTGAACGACGCAAAACCGGCTGCGGCTTATGCTAT 4411  
 Qy 4701 CAGGATCAGGCTTTACTGTTGATGCGGCTGCGGATTAACCGATCCAGGCTATC 4760  
 Db 4412 GAATATCAGGATTCATTTGATGACGAGGCTTACCGTTATACCGATCCAGGCGCAAT 4471  
 Qy 4761 GAGCGCTGTTTACGCTGCGGAGCAAGCACTCATGATTAATGCTGACCTGATCCGCTG 4820  
 Db 4472 GAABAATCTTTACCTGCGGCGGAAAACGTTAAAGATTACGTTGAACTGCTCCGCT 4531  
 Qy 4821 ACGCCATTTTATGCGCTGCTGCGGAAAGACGCGAGCGCTGACATACGACAAATGAG 4880  
 Db 4532 GCGCGCTTTTATGCGCTGCTGCTGGAAGTCAAGAAAGTTTCAATTAAGAAATGATCAG 4591  
 Qy 4881 GCGGAGCTGAGAGAGATTGCACTTTTAATCCCGAGATGTCGCGGCTTACCGGCGAG 4940  
 Db 4592 GCACAGCTTGAAGGCGGAGATTCAGCAGATTATACCGGATGTTGAAGGCTATGCTAG 4651  
 Qy 4941 TTTCTGCGCTATTCACAGATGCTTTTCTGTAAGGCTATCTGAATGCGGACCGTACT 5000  
 Db 4652 TTTCTGCGCTATTCACGCGGCTGTTTAAAGAAAGCTATCTGAAGCTTGGACGCTGCT 4711  
 Qy 5001 TTTCTGCGCTTTCGCGAGATGCTGCTGCGGCGCAACGCTGCGCTGCGGCTGCG 5060  
 Db 4712 TTTCTGCGCTTTCAGGAGATGCTTTCGCGCGCGCCCACTGCGGAAACTGCAAGCATG 4771

Qy 5061 CGGATGCTTACAGCATGCTGCGGAAATTTATTCATGACATCATCTGCGCAGGCTTT 5120  
 Db 4772 CGAACCGTTTACAGTAAGTAAGTTGCGAGCTACATTAAGGATATAATCTTCTCGGCGCTT 4831  
 Qy 5121 TCCCTTACTCGTTGCGGATCGGCGGATATCTTTTTCGCAAGCTTCGATCTATACCTTA 5180  
 Db 4832 TCTTTTACTCGCTGTTAGTGGGCGGCAATCCCTTGCCACCTCTCTCATTTATACCTTA 4891  
 Qy 5181 ATTACGCACTGAGACGCGAATGCGGCTGCTGCTTCCGCGCGGCTACCGGCTGCGCTG 5240  
 Db 4892 ATACATGCTATCGAAGCGAATGCGGCTGCTGCTTTCAGCGCGGCTGTAACGCGGCACTTA 4951  
 Qy 5241 GTTATGAGCATGCGCGGCTGTTTTCGGAATTTGGCGGATTAATGCTGCTCAACCGGAA 5300  
 Db 4952 GTGAAAGGATGATTAAGCTGTTTCAAGATCTGGTGGCCAAAGTGTACTGAATGCGAG 5011  
 Qy 5301 GTCAGCAGCTGAGACCGAGGTTAACCGCATCAGCGGCTGCAAGCTGAAGGATGCGCG 5360  
 Db 5012 GTACGCGCATGGAACCAAGGCGGATACCATTAAGACCGTGCATTTAGAGACGAGCG 5071  
 Qy 5361 CGTTTCCGCGCGCGCGCTTGCCTCAATGCTGACGCTGATCATCATGCTGCTG 5420  
 Db 5072 AGATTCGCGACCGGCGCTGCTGCTCAATGCGAATGCTTCAACATTCGCGACTG 5131  
 Qy 5421 TTAAGCCAGCATCTGCGGCGGTTAAACGCGCGCAACGCTGAAGCCGAGCGGATGAG 5480  
 Db 5132 TTAAGTCAGCATCTGCGCGCTGAAGCAGTCAAAAACCTGCAAACTGAACGCGATGAG 5191  
 Qy 5481 AACTGCTGTTTACTATCTATTTTGTCTTAATCATGCGCACCGGAGCTGCGGCAAC 5540  
 Db 5192 AACTCATTTTGTCTCTATTTTGTCTGATCATCATCATCATCATCATCATCATCATCAT 5251  
 Qy 5541 ACAGTGTCTTTGCTGCGGCTATCTGTAATGATGATGATGATGATGATGATGATGATGAT 5600  
 Db 5252 ACAGTGTCTTTGCTGCGGCTATCTGTAATGATGATGATGATGATGATGATGATGATGAT 5311  
 Qy 5601 CTGCGGAGAAATTTCTGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 5660  
 Db 5312 CTGCGAGCAATTTCTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACT 5371  
 Qy 5661 CCGCGGCTGCGGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5720  
 Db 5372 CCGGAAAGTTGCGGAGTTACTATGTTAGTCTCCGCTGCAATTAAGCAAGCGAAC 5431  
 Qy 5721 ATTGACTGCAACAGAAAGGCGCGCTTGGCGGATGCAATTTTGTCTTCTGAGAGAG 5780  
 Db 5432 CTGCACTGGAACGATGGAAGGCGCGCATGCGGATCTGATTTTGAATGATGATGATGATGAT 5491  
 Qy 5781 CACTATGCTGCGGCTGCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5840  
 Db 5492 CATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5551  
 Qy 5841 TTTTGCAGACGCTGCAACGCGCATCAAGGCTCAAGCTTTTGTGCTGCAACCATTTTGAAG 5900  
 Db 5552 TTTTGCAGACGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCTCAATGCT 5611  
 Qy 5901 CAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5960  
 Db 5612 CAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5671  
 Qy 5961 GGGGCTGCGGATCCCGGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 6020  
 Db 5672 GGTGCAAGCAACCATCCCGGCGGCGGCTTCCGCGGCTGATCCGCTGCTGCTGCTGCTGCT 5731  
 Qy 6021 GCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6080  
 Db 5732 GCAGGTTTATGCTGAGGAGCTGA---TTGAAATATCGTGTGTTACTCAATCATGCGG 5788  
 Qy 6081 CGAAACATGCGGCGAGGCTCCAAAGTTTCCGCAAGCTTACCAGGCTATTTGATCTT 6140  
 Db 5789 TCGAAACGATGCTGTTGCTGCAAAAATTTTGCAGCTGCTCAAAATTTTGAACGCA 5848



[illegible][illegible]



OY	2196	AAAGCGCTGCGCGCGGGGACATGGACATCACCTTATCCAGCAAAAGTATGACAGCACT	2255
Db	5628	TGGAGATGGCCCAACGGCGCCACCGGGGACCTTTCTC-----ACGGACACGTCGCCT	5681
OY	2256	TGCTGAGCGACGACGCGCATTTGCTTTTGGCCGTGGGAGCGGACCGCATCTCCGGAT	2315
Db	5682	CGCTGGCAGAGCAAGAAACGGAGCGGGGTGGTCTATCCACTTCCCGCAGCGTGCAC	5741
OY	2316	CGCTTCCAGGAACTCAGCGCGCTGGCCGACCGGGCGGGCTGTGCTGTTGACGTGA	2375
Db	5742	AGGCCCAAGCGCAAGTCCAGCAGAGAGTAAACGGCAACTCTCGCGGCTGATTGCGGCA	5801
OY	2376	TTCAACGATCTTGGCAGACCAACCAATATGCTATACCGCGAACCTCCCGCGGTGCTGCAC	2435
Db	5802	TGTCACTCCCTA-----ACCAATGTCCTGTGCACAGATTTCCGCTATTTCTAACGC	5852
OY	2436	GGCTGCAGAGTCCGATGGCGTGAATGGCCGATCAATAGAAAGGAGTGTGATCTGTGGCAG	2495
Db	5853	GGCTGGCGGTGGAACCGCTGATTTGTCCATGAGATGAGAGCCGCGCGAAAGCTGGATTGCGC	5912
OY	2496	AGGCGTTACAGCTGCCGTTGCTGTGTCGTGGTGCCTGCGCGCTGCGCGGTCAATGCGAAGCG	2555
Db	5913	AGGCGCTGAGCACTAACATTTATCTCTATTTGCTGTGGCGCTGCGGTCACCGCAGCGGG	5972
OY	2556	CCATTTCCGCTGTGTGTGATGCTCCCTTTGCTTTGCTTGTCTCAGATGAGAAAAGCGTGCAGCGCT	2615
Db	5973	GTCGTGCGGCTCCGAGTGAATGCGGTTTCACTAGCGCGAGAGTAAAGAGCCCTGGGGGCTT	6032
OY	2616	ATCAGAGCCAGCAGATGACATCTACGACCGCATCATGCGTGTGATAGGCGCTGTCAATCGCTC	2675
Db	6033	TTCAAGTTCAGGAAACGATCTTACATGCGCTGATGTACCCGACGGGACAGACGATCTGCG	6092
OY	2676	GTCATGCGCGCGCTTGGCCTGCGCCGAAACGCAATGCTTACATCATGTGTCGTGCGCGC	2735
Db	6093	GCCACGCCCAACGCTTGTGTTTGGCGGAGCGCAGGCGGTCTGACGAGTGTCTTCCGCGC	6152
OY	2736	TGCGGCGCAATCAGTACGCTGTGTCGCTTTTGAATTTTCCACGCGCAGCAATGCGCAGCT	2795
Db	6153	TGGGCGAGATTAAAGCACATGCCGTTCCGGCCCTGCAGCTTCCACGCGCGGCGCTGCCAACT	6212
OY	2796	GCTATCACAGCGTGGAGTCCGCTGCGGGAATCTCAGATTGCTAGCGGCGGCTGCGCCACCT	2855
Db	6213	GTTTTCACATCAGTGGAGACCTGGCGTATC-----AGCCGCGCGCAGGTAGAAC	6263
OY	2856	GGCCACGCGCTGCGCCAGCGCGGTGTATATGCTGCTGGGACGCTACAGGGGATCGCT	2915
Db	6264	GCTGCGCACGACGACCGCGCGGATCTTTGCTGCTGTGGGACCTCCAGGGCCACGCTC	6323
OY	2916	TTGCGCTGTTTCTGCATATGTGCTCAGGCTGACGCTGCGCAATCAGACGCTGTGCTGGTG	2975
Db	6324	TACCCCTGTATTCAAGAAAGATCGCCGCGCTGTGCGAGGTGGGGGCGGAGGTGACCATTG	6383
OY	2976	CACACTGTGGGGGTGACCGCCGACGACGAGACATAGCTCAGACTGGCGCGGTCTGCT	3035
Db	6384	CCCACTGAGATGGCTGACGCCCGCCGACAGGCGAGATGCTCTTAAGCTGCGCGGCGAGCG	6443
OY	3036	GGGATGACCGATTTTGTGATCAGGCGGCGGCGCTGACAGCATGCGCAACTGTTTATCACTC	3095
Db	6444	AGGTGTGATGCTTTGTGTGACACAGCCGCGCTATGTTGCCAGGGCTAATGTGATATCAACC	6503
OY	3096	ACGCCGAGTCTGAACAGTGGCTGGAAGACTGGAAGTGTGACACGCGCGATGCTGGCGCTGC	3155
Db	6504	ACGGCGAGTCTCAATACCGTACTGATGCGCTGGCTGCGCCGACCGCGGCTGCTGGCGGTGC	6563
OY	3156	CGATTCGCTTGTGATCAGCGCGGCGGTGGGACGATTTGAATGGGACGGCGCTCGCGCGCC	3215
Db	6564	CACCTCTTTTGACACGCGCGCGGTGGCTGCGCCGCGGTGTCTAATACGGGCTGTGGTGC	6623
OY	3216	GGCGCTCAGCTTTCAGCGCGGTGCGGACGCTGAGACCAACTGCAACATGTGCTGATGTG	3275
Db	6624	GGGATAGCGCTTTTGGCCAGACAGACGAGCGCTGGGGAGATTTGGCCCACTGCGTGGGG	6683
OY	3276	ACGATGCGTATGCTGTGCGCATGTTCAGCATTCAGGCGCACTGACGCGGCGGATGCT	3335

Db	6684	ATGAGACGCTGCATCACCCTCTCGCCACCGCCGCCACAGCTTAACAGCGCCGG66CA	6743
OY	3336	GTAAGCGCGCGGCTGATATTTGTCAGACAGCGGCTGTGTACAGCAAAATCGTCTGCGCG	3395B
Db	6744	CGCCCGGCGGACCCCTGATTGAACAGGCCATAGACAGGAGTGAAGAGCGTATCTGACG	6803A
OY	3396	AGGCACTGATGCGCGCACCTTATGATGTCTTCTGTGGTGC CGGCTTGGTTACCG	3455
Db	6804	-----GGAATCGATTTTATGTGGCGGGCGGCTTGGCAACGG	6839
OY	3456	GCTGATTTGGCGTGGATTAAAGCCAGCTGCAGCCCGCACTTAAGGTTTGTACTGAGAG	3515
Db	6840	GCTATGCTCGGCGGTCTGGCGCACGGCTACCCGAGCTTAACCTGCTGCTGATCGAGC	6899B
OY	3516	TCAGCGCAGCGCGGCGCGCAATCAATACCTGTGTCTTCATCGCGAAGACGTCAAGC	3575
Db	6900	CGGGAGAGAGCCCGCGGGAACATACCTGGTCAATTCATGAAGACGATCGATCCCGG	6959B
OY	3576	GCACTTTCGCTGGCTCGAGCGCGCTCTTTGGGCGCGCTGGCCCGGTTATAGATACGCTT	3635
Db	6960	GCACACAGCCCTGGCTGGCCCGCGCTGGGCCACCGCTGGCGCGGCTATGAGGTGACGTT	7019
OY	3636	CCCCACCCTGGTGGCGCAGCTGGAATGGGAATATTTGCTCATTTCTCTCGAGATTTTGC	3695B
Db	7020	TCCGATTTTGGCGTGGCTTCCGCGCGGCTTACTTCTCATTTACTTCAGAGCGCTTTGC	7079
OY	3696	CCGCACTTACAGAGGTGTCTCGGTGCGGCTTAACGACCGCAGCGCGCGGTCAAGGT	3755
Db	7080	CGAGCGCTGATCAGGCGGCTGGGGGAGAACATCTGGCTAAACGTTCGGTGAACGAGGT	7139
OY	3756	CTCACCCACCGGGGTCAAGCTGCGCGGATTTGCGGGAATTTCAGCGCGAGGCGGTGATTGA	3815
Db	7140	GTTTACCCATATGCTGTGGCTTTTGGCAACGATGAGCGCGCTGCTTTCGGAAGGCGGATTTGA	7199
OY	3816	CGGACGGGGCGTGAAGCGCAACCGGCATCTGAGCTGGGCTATAGGCAATTTGTCGTC	3875
Db	7200	CGAGAGCGGCGGTGACCGGCACTTGGGAGTGAACCGGCTTATAGCTCTTTCTTGGTCA	7259
OY	3876	GGAGTGGCAATGCGCGCGCGCATGCGCTTGCACAGCAACATATTTGATGACGCGCAGGT	3935
Db	7260	GCAATGGGCGGTGACACAGCGCCACGGCGCTGACGTAACGATCTGTATGATGCGACGGT	7319
OY	3936	CGATCAGACAGGGGTTATGCTTTGTTTACACCTTGGCGCTCACTGTCAGCGCTTTACT	3995B
Db	7320	GGCCACAGAGAGGGCTATCGCTTTGTCTTACACGCTGCCGCTCTCCGCGCACAGCTGCT	7379
OY	3996	GATTGAAGATACCCACTATCAATCAATGCCAGCGCTGATATGCCGACAGCGCGCCGCTCA	4055
Db	7380	GATGAGAGATACGGCTACGCAATGTGCCGACCGGTATGATATGCTTACGCGCAGAC	7439
OY	4056	CATTACGATATATCCACACAGCGCGGCTGGAAATTTGCGCAGCTGTGCGCGAGAGCA	4115
Db	7440	GATTACGACTATGCTACAGCAAAAGGTGCGACTGGCCAGCTTGAACCGCAGAGAC	7499B
OY	4116	CGGCTCGTGGCATCAAGCTCAAGCGCGGATATGATCACTTTGCGCAACGACAGCAG	4175
Db	7500	CGGCTGTCTGCGATTAACCTTGGCGGGGTGACATCAAGGCTCTGTGGCGCGATGCGCGG	7559
OY	4176	GCAACCGTGCAGCGGGCTGGCGCGCGGACTGTTTTCAGCGCACCAACCGTTACTGGTGC	4235
Db	7560	CGTCCCGGCTCGGGAATGCGGGCTGGGCTATTTTACACCTTACACTGCGTATTTGCTG	7619
OY	4236	CGCGCGGTGGCGGTGCGGAGAAATTTGCAGACAGCTGGCCCGCGCACGCTCACAGCT	4295B
Db	7620	CGTGGCGGTGGCTTGGCCAGCGGATTTGCCAGACGCGCGCGCTGGGCAAGCTTCCGCT	7679
OY	4296	GAGCACTGCATCGAATCTCTTGGCGGTCAAGCTGCGCGGAGACAGCGCTTTTCTGCTT	4355
Db	7680	CTATCAGCTCACCGGAGGTTTGGCCGAACGCACTGGCGCGAGCGAGGATTTCTTCCGCT	7739
OY	4356	GTTAAATGCAATGCTGTTTCTTGGCGGACAGGCTTGAAACGCGCTGGGCGGTAATGCAAGC	4415

Dh 7740 GCTGAACCGGATCTTTTCTGCGCGGCGGAGAGAACCGCTGCGGGTATGACGC 7799  
Qy 4416 TTTTACCGGTTACGCGCGATGATTAAGCGCTTTTACCGCGGCACTGCGCCTCAG 4475  
Dh 7800 CTTTATAGGGCTGCGGAGCCACCGTAGAGCGCTTTTACCGCGGCTCTCTCTT 7859  
Qy 4476 CGATAAAGCAGCAATTCGTGTCGGCAAAACCGCGGTCCCTCTCGGCGAAGCGCTGCGCG 4535  
Dh 7860 TGAATAAGCCCGCATTTTGAAGGGCAACCAACCGGTTCCGTGGCGAAGCTTGCGGCG 7919  
Qy 4536 ATTGATGATGACCTCTCGTTTACAGAGAGAAATAATGAACCGACCTATGATTTGCG 4595  
Dh 7920 GGCCTGAACCATTTTCTGACAGACAGATTAAGATGAATAAAACCGTGTGATTTGCG 7979  
Qy 4596 GCAAGCTTCGTTGCGCTGCGCTGCGCTGCGATTTGTTGCAAGCGCGCGCGCTGCGGTACG 4655  
Dh 7980 GCAAGCTTTGGTGGCTGCGCTGCGCTGCGATTTGCGCTGCGCGGAGGAGATCCCAACGTA 8039  
Qy 4656 CTGCTGGAACAGCCGATTAAGCTGCGCGCGCGCTTATGATTCAGATCAAGGTTTTT 4715  
Dh 8040 CTGCTGAGAGCAGCGGAGACAGCCCGCGGTGCGGCTTACGTCTGCGCATGACAGGCTTT 8099  
Qy 4716 ACCTTTGATGCGGTCGCGAGGTATTAACCGATCCAGCGCTTACAGCGCTGTTACG 4775  
Dh 8100 ACCTTTGACGCGCGCGCGAGCGGTATCAACGATCTTACCGGCTTGAGGCGCTGTACCC 8159  
Qy 4776 CTGCGAGGCAAGCACTCAGTATTAATGATGATCTGATTCGCGTGAAGCGCATTTTATGCG 4835  
Dh 8160 CTGCGCGGAGCGCATGAGAGATTAAGTCAAGCTGCTGCGCGTAACCCCTTCTACCGA 8219  
Qy 4836 CTGCTGAGGAAGACGCGAGGAGCTGAGCTGAGCTACAGCAACATCAGGCGAGCTGAGAG 4895  
Dh 8220 CTCTGCTGAGAGTCGCGGAGAACCTTCGATCTATGCTACAGACGCGAGCTTGAGGCG 8279  
Qy 4896 CAGATGCCACTTTTAAATCCCGAGATGTCGCGGTTACCGCGAGTTTCTGCGCTATTCA 4955  
Dh 8280 CAGATTAACCGAGTTCAACCCCGCGAGCGTCAAGGAGGATACCGGCTTTCTGCTTACTCC 8339  
Qy 4956 CAGATGTTTTCGTGAGGGCTATCTGAACCTGAGGACCGTACTTCTTCTGCACTTCCGC 5015  
Dh 8340 CAGGCGGATTCAGAGAGGATTTTTCGCTGCGAGCGGTCTCTCTCTTCTGCG 8399  
Qy 5016 GACATGCTGCGTGCAGGCGCACAGCTGGGTGCGCTGAGGCGCTGCGGAGTGTCTACAGC 5075  
Dh 8400 GACATGCTGCGCGCGCGCGGCGGAGCTGTTAAGCTTCAAGGCGTGGCAGAGGTCTACAG 8459  
Qy 5076 ATGATGCGAAATTTATTCATGACGATCATCTGCGCAAGGCTTTTCTCTCTGCTG 5135  
Dh 8460 TCGGTTTCGCGCTTTATTTAGAGATGAGCATCTGCGGAGGCTTCTGCTTCACTCCCTG 8519  
Qy 5136 CTGCTCGCGCGTATCTTTTGAACGCTTTCGATCTTAACTTAAATTTCAAGCACTGAG 5195  
Dh 8520 CTGATAGCGGCAACCCCTTCAACACCTGTCATCTAACACCTGATTCAGGCGCTTGAG 8579  
Qy 5196 CGCAATAGGCGCGTGGTTTCCGCGCGGCGGATCCGCTGCGCTGATGATGAGCGG 5255  
Dh 8580 CGGAGTGGGGGGTCTGCTTCCCTGAGGGCGGCAACCGGCGCTGCTGAGAGCGATGCTG 8639  
Qy 5256 CGGCTGTTTTCGCAATTTGGGCGGTAACCTGCTCTCAACCGCGGAGTCAAGCAGCTGAG 5315  
Dh 8640 AAGCTGTTTACCGATCTGAGGGGGGAGATCGAATCTCAACCGCGGTCGAAAGCTGCTG 8699  
Qy 5316 ACCGAGGATACCGCATCAGCGGTGTCCAGTGAAGATGGGCGCGCTTTTGGCGCGCC 5375  
Dh 8700 GTGCGCATTAACCGCGTAGAGCAGGTCCGCTGCGGATGCTCGATCTTTGACACCGAC 8759  
Qy 5376 GCGCTGCGTCAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5435  
Dh 8760 GCGCTAGCTGAGCGCTGAGAGTGTGAACCTTAATAAAAGCTGCTGCGGCACTATCCG 8819  
Qy 5436 GCGGCGGCTAAACCGCGGCAACCGCTGAAGCGGAGGAGTGAACCTGCTGTTTGTGA 5495  
Dh 8820 GTGGGGCAGAAAGCGGGCGGAGCGCTGAGGCGAAAGCATGACCACTGCTTTGTG 8879

Qy 5496 CTCATTTTGGTCTTAATCATGCCAACCGGACAGTGGCGCACACACGATGCTTTGT 5555  
Dh 8880 CTCATCTTGGCTGAAACGACCTCATTTCCAGTGGCGCACCAATCATCTTTTGT 8939  
Qy 5556 CCGCGCTATCGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5615  
Dh 8940 CCGCGCTACCGGAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8999  
Qy 5616 TCGCTGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5675  
Dh 9000 TCGCTACCTGCACTCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9059  
Qy 5676 AGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5735  
Dh 9060 AGCTTCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9119  
Qy 5736 GAAGGCGCGCTTGGCGGATGCGCATTTTCTTATCTGAGAGAGCATATATGCGGCT 5795  
Dh 9120 GAGGGGCGAAGCTGGCGGACCGCATCTTTGATCTTGAAGAGGCTATATGCGCGC 9179  
Qy 5796 CTGCGACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5855  
Dh 9180 CTGCGATGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9239  
Qy 5856 CAGCGCATGAGGAGCTGAGGCTTTGCTGAAACCATTTTGAAGCAAGCGCTGCTG 5915  
Dh 9240 GATGCGATCTGAGATGAGCTTCTCTCATGAGCGCTGCTGACCAAGAGCTGCTG 9299  
Qy 5916 CCGCGCATACCGCGATGCGGATTCATTTATCTTATCTGATGAGGCGTGGCAGCAT 5975  
Dh 9300 CGCGCGCAACCGCGGACGATTCGCACTTTCATGAGCGCTGCTGACCAAGAGCTGCTG 9359  
Qy 5976 CCGGTCGCGGTGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6035  
Dh 9360 CCGTGGGCGGAGATCTCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9419  
Qy 6036 GAGGATGACCGGATGAAACCAACCGCGCTGATGATGATGATGATGATGATGATGATGATGAT 6095  
Dh 9420 GAGGATG--CAATGAGCAACCGCGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 9476  
Qy 6096 AGGCTTCGAAGATTTGCGCGAGCTAACCGGCTATTTATCTTCAACCGCGCGAGTA 6155  
Dh 9477 ACGGCTGAAAGATTTTGGCCACCGCTGCGAAGCTGTTGACCGCGGCAACCGCGCTGAGCG 9536  
Qy 6156 CGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6215  
Dh 9537 TGCATGATCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9596  
Qy 6216 GCGAAGCGGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 6275  
Dh 9597 GCTTTCGCAAGAGCGCGCGGAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 9656  
Qy 6276 TCGAAAACCGCGCGCTTACAGCGGTCGCCACATGATGAACAGAGCTTTCTGCTTTC 6335  
Dh 9657 CGCTGACCTTGGGGGGTTTGAAGGGGCGAATGACAGATCCGGCTTTCGCTGCTTTC 9716  
Qy 6336 AGGATGCGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 6395  
Dh 9717 AGGAGTGGGCGTGAACCAAGATTAACCGCGGAGTGGGCTGATGATGATGATGATGATGATGAT 9776  
Qy 6396 TTCGAT 6455  
Dh 9777 TTTGAT 9836  
Qy 6456 ATACGATGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6515  
Dh 9837 ATACGATGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9896  
Qy 6516 GCGTACTGATCAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6575  
Dh 9897 GGGTGTGAT 9956



```
QY 3523 CAGCCGCGCCGCAATCATCTGAGTCGTTCCATGCGGAAGACGTACAGGAAGCCGAGTT 3582
Db 245 CAGCGCGCGGGAAATCATACGTGTCATTTTACACAGATGATTTGACTGAAGCCAAAT 304
QY 3583 CGCTGCGCTCGAGCGCGCTGCTTTCCGCGCGCTGAGCCCGGATTCAGAGTACGTTCCCGAC 3642
Db 305 CGTTGAGTAGTCGCGCTGCTGCTGTTTCATCATCTGCGCCGACATACAGTACGTTTCCACA 364
QY 3643 CTGCGTGCAGCTGAGTGTGAATATTGCTCATGCTGCTGCGAGATTTTGGCCGCGAC 3702
Db 365 CGCGGTGATGCTGAACAGCGGCTACTTTGTATTACTTCTCAGCGCTTTCGCTGAGGTT 424
QY 3703 TTACAGAGGTGCTCGGTGCGCGCTACGACCGCAGCGCGGTGACGAGGTCTCACCC 3762
Db 425 TTACAGCAGCATGTTGGCCCGACTTGTGATGATACCGCGGTGCGAGAGTTAATGCG 484
QY 3763 ACCGCGGCTCAGACTGCGCGATGCGCGGATGTTACAGCGCGAGCGGTGATGACGACG 3822
Db 485 GAATCTGTTGCGTTGAAAAAGGTCAGGTTATCGGTGCCCGCGGTGATTTGACGCGCG 544
QY 3823 GGGCTGCAAGCCGACACCGCATCTGCAAGCTCGGCTATACGCAATTTGCTGAGAGTGG 3882
Db 545 GGTATGCGGCAATTTCAGACTGACGCTGGCTTCCAGCGTTTATTTGGCCAGGAAATGG 604
QY 3883 CAATCGCGCGCGCATGCGCTGCAAGCAGCAATATTGATGACGCGCAGGTGATTCAG 3942
Db 605 CGATTGAGCCACCGCGCATGTTTATCGTCTCCATTCATGATGACCGGTGATTCAG 664
QY 3943 CAGCAGGTTATCGCTTGTGTTACACCTGCGCTCAGTGCAGCGCTTTTACGATTGAA 4002
Db 665 CAATATGTTATCGCTTGTGTTACAGCTGCGCTGCGCGACCAAGTTGTTAATTGAA 724
QY 4003 GATACCCATCATACATACCATATGCAACGCTGATGCGGACAGGCGCGCTGACATTAAC 4062
Db 725 GACACGCACTATATGATTAATGACACATTAATCTGATGCGCGCGCAAAATTTGTC 784
QY 4063 GATTATGCCACACGCGCGCTGGAATTTGCGCAGCTGCTGCGCGAGGACAGCGGCTCG 4122
Db 785 GACTATGCGCGCAACAGGTTGGAGCTTACAGACCTGCTGCGAAGAAAGAGGCGCC 844
QY 4123 CTGCGCATCAAGCTCAGCGCGGATATGATCATGTTCTGGCAACAGACAGGCGCAACCG 4182
Db 845 TTACCATATTCGTGCGGCAATGCGGACGCAATTCG---CAGCAGCGCGCCCTGCGC 901
QY 4183 TGCAGCGGCTGCGCGCTGCACTGTTTACCGCACACCGGTTACTGCTGCGCGCGCG 4242
Db 902 TGTATGATTAATGTCGCGGTCTGTTCCATCTTCAACCGGCTATTCATGCGCGCTGCG 961
QY 4243 GTGCGCTGCGGAGAAATGTCAGCACTGCGCGCGCGCACTGACACGCTGAGCAC 4302
Db 962 GTTGCCTGCGCGACCGCTGAGTGCATTTGTTACGTGCGCTCAATTCACCAT 1021
QY 4303 TGCATCGAATCCTTTGCGCGTACGACCTGCGCGAGCAGCGCTTTTCCGTCTTAAT 4362
Db 1022 GCATTTAGCATTTTTCGCCGAGGCGTGCAGACGAGGCGCTTTTCCGATCTGAT 1081
QY 4363 CGCATGCTGTTTCTTGGCGGACGCGCTGAAACAGGCTGCGGTAAATGCAAGCTTTTAC 4422
Db 1082 CGCATGCTGTTTAAAGCGGACCGCGCATTCACGCTGCGGTAAATGCAAGCTTTTAT 1141
QY 4423 CGGCTTGA CGCGGATGATTAAGCGCTTTTACGCGCGGCAACTGCGCTCAGCGATTA 4482
Db 1142 GGTTTACTGGAAGATTAATTTGCGCGTTTAAATGCGGAAATTCACGCTTACCGATCG 1201
QY 4483 GCAAGCATTTCTGTGCGGCAACCGCGGTCCCTCTCGCGGAAGCGCTGCGGCAATTGAG 4542
Db 1202 CTACGTATTTGAGGCGGCAACCGCGTGTTCGCGATTAAGCAGCATTTGCAAGCATTAATG 1261
QY 4543 ATGACCTCTCGGTTACAGGGAAGAAATTAAGAAACGACCTTAATGATTTGGCGAGGCT 4602
Db 1262 ACGACTCATTCGTTAAAGGACACTA---CATGAAACCAACTACGCTAATTTGGTCAGGCT 1318
QY 4603 TCGGTGCGCTGCGCGTGCATTCGTCTGCAAGCGGCGGCGTCACGGTCACGTCTGCG 4662

Db 1319 TCGGTGCGCTGCGCACTGCGCAATTCGTCTACAACTGCGGGGATCCCGCTCTTACTGTTG 1378
QY 4663 AACAGCGCATTAACCTGCGCGGCGCGCTATGTGATACAGATTCAGAGTTTAACTTTG 4722
Db 1379 AACAACTGATTAACCCGCGGCTGCGGCTTATGTCTACGAGATTCAGAGGTTTAACTTTG 1438
QY 4723 ATGCGGCTCCGAGCGGTATTAACGATCCGACGCTATCCGAGCGGCTTTAACCTGCGAG 4782
Db 1439 ATGAGGCGCCGAGCGGTATACAGGATCCAGTCCAGTGCATTAAGAACTGTTTGCAGT 1498
QY 4783 GCAAGCACTCAGTATTAATGCACTGATGCGGAGCGGATTTTATTCGCTGCT 4842
Db 1499 GAAACAGTTAAAGATATGTGAATGCTGCGGTTATGCGCGTTTACCGCTGTGTT 1558
QY 4843 GGGAAAGCGCAGGCACTGGACTACGACAATATCAGGCGCAGCTGAGACAGATTCG 4902
Db 1559 GGGAGTCAGGGAAGGTCTTTAATTACATACATCAAAACCGGCTGGAAGCCAGATTC 1618
QY 4903 CCACTTTAATCCCAAGATGTGCGCGTTACCGGCACTTTGCGCTTATTCAGATG 4962
Db 1619 AGCAGTTAATCCCGGATGCGAAGTTATGCTAGTTTCTGAGACTATTCACGCGG 1678
QY 4963 TGTTCGTAGGCTATCTGAAAATGCGGACCGTACTTTTCTGCAATTTCCGCGACATTC 5022
Db 1679 TGTTTAAAGAAAGCTATCTAAAGCTGAGTACTGCTTTTATATGCTTACAGACATTC 1738
QY 5023 TCGCTGCGGCGCAACAGCTGCGTGCCTGACAGGCTGCGCAGTGTCTACATGAGTGG 5082
Db 1739 TTGCGCGCGCACCTCAACTGCGCAAACTGACAGCATGAAACGTTTACATGAAAGTTG 1798
QY 5083 CGAAATTTATTCATGACATCTGCGCAGGCTTTTCTTCACTGCTGCTGCTG 5142
Db 1799 CGAGTATCATGAAAGATGAAATCTGCGCAGGCGTTTCTTCACTGCTGCTGCTG 1858
QY 5143 GCGGTATCCTTTTGCACACGCTCTGCACTTATCTTAACTTTCACGACTGAGCGCAAT 5202
Db 1859 GCGGCAATCCCTTCGCGACCTCATCACTTATATGTTGATACAGCGCTGAGAGTGA 1918
QY 5203 GGGCGGTGTGTTTCCGCGCGCGGATACCGGTGCGTGTGATGCGACGCGGCTGT 5262
Db 1919 GGGCGGTGTGTTTCCGCGCGCGGACCGGCGCATTAAGTTACGGGATATTAAGCTGT 1978
QY 5263 TTGCGCATTTGGCGGATGCTGCTGCAACCGCAAGTCAACGCTGGAAGCGAGG 5322
Db 1979 TTGAGATCTGGGATGGAAGTGTGTTAAACGCAAGTCAACATATGAAACGACG 2038
QY 5323 GTAACCGCATCAGCGGTGTCAGCTGAAGATGAGCGCGCTTTTGC CGCGCGCTTG 5382
Db 2039 GAAACAAAGATTTGAAGCGGTGCACTTTAGAGGACGCTGCAAGTTCGACCAAGCGTGC 2098
QY 5383 GGTCAATGCTGACGTGTGCAATCTTACGATGCGCTTTTAAAGCAGCATCTGCGCGC 5442
Db 2099 GGTCAATGCTGATGATGTTTCAATCTTATGCGACCTGTTAAACGCAAGCTGCGCGC 2158
QY 5443 GTAAAGCGCGGCAAGCTGAAGCGCAAGGAGTGAAGCACTGCGTTTGTATCTAT 5502
Db 2159 TTAAAGCTCAACAACTGCAAGCTGAAGGCAATGATCTCTGTTGTGCTTAAT 2218
QY 5503 TTGCTCTTAATCAATGCCACCGGAGCTGCGGACCAACAGGTGTGCTTTCGCGCT 5562
Db 2219 TTGCTTTGAATCAACATCAATGATCACTGCGGATCAACAGGTTTGTTCGCGCGGCT 2278
QY 5563 ATGTGAATTAATGATGATGATCTTCAATAGCAAGCAGCTGCGGAAAGATTTCTGCTGT 5622
Db 2279 ACCGCGAGCTGATTAAGAAATTTTATCATGATGCTGCGGAGGAGCATCTTCACTTT 2338
QY 5623 ATCTGATGCGGCTGCTGCGAGATCCGTGCTGCGACCGGCGGCTGCGGAGTTT 5682
Db 2339 ATCTGACGCGGCTGCTGCGAGATCTGATCTGCGGCTGGAAGGTTGCGGAGTTACT 2398
QY 5683 ACGTCTGCGCGGCTGCGCATCTGCTGATCCCGCAATTTGACTGCGCAACAGAAAGGC 5742
```

Db	2399	ATGTGTTGGCGCCCGATGCGCATTTAAGGCA	CCGGAAACCTCGACTGACGGATTGAGGGGC	24548
QY	5743	CGCGCTTGGCGGATTCGCAATTTTGTGCTTA	CTCTGGAGAGACCTAATATGCCGGGCTCTGGCAC	5802
Db	2459	CAAAACTAACGGGACCGGATTTTTTGTGGTAC	CTTGGAGCAGCACTTAATGCTGGCTTACCGGA	2518
QY	5803	AGCAGTAGTAGACACAACCGTAATGTTTAC	GGCGGTTGAATTTTCGGCAACAGCTGCACGGCGC	5862
Db	2519	GTCAGCTGGGTACCGACCGGATGTTTACCG	CGGTTGATTTTCGGACCAAGCTTAATGCTC	2578
QY	5863	ATCAGGGCTCAGCGTTTYTGCTCGAACCCAT	TTTGTGACCGAAAGCGGCTGTTCGGCGCGC	5922
Db	2579	ATCATGGCTCAGCGCTTTTCTGTGGAGCCG	TTCTTACCAGAGGCGCTGGTTTGGCGCGC	2638
QY	5923	ATTAACCGGATTCGCGACATTTACTTAAC	CTTTATCTGTGTGGGAGCTTGGCACGCAATCCGGGT	5982
Db	2639	ATTAACCGGATTAACCACTTAATCTCTAC	CTGTGTGGGCGCAGGCGATCCCGGCG	2698
QY	5983	CCGGATGTCGAGGGGTGATTCGGGCTCCG	GGAAGGGAACCGCCACGCTGATGTGGAGGATC	6042
Db	2699	CAGGCAATTCCTGGGGTCAATCGGCTCG	GAAAGGCAACGAGGTTGATGTGGAGGATC	2758
QY	6043	TGACCGGATGTAACCAACCGCGCTGATTT	GAGCAGGTCAACGAAACATATGGCGCAGGGCTC	6102
Db	2759	TGATTTGAATATTCGTCGTACTCAATCA	TCGCGCTC---GAAACGATGGCAGTTGGCTC	2815
QY	6103	CAAAAGTTTGGCCAGCGCTACCCGGCTAT	TTTGATCTTCAACCGCGCGCAGTACGCTGAT	6162
Db	2816	GAAAGTTTTTTCGACAGCGCTCAAAAGTT	ATTTGATGTCAAAAACCGGCGCACGCTGAT	2875
QY	6163	GCTGTACGCGCTGTGTGTGTGTCACTGT	GAAGATGTGATAATGTGACAGCGTGGGGGAAG	6222
Db	2876	GCTCTACGCGCTGTGTGTGTGTGTGTG	AGATGTGATAATGTGACAGCGTGGGGCTTCA	2935
QY	6223	CGGACGAGGACGCGGTGTGTGTGTGTG	ACACGAGCGGATGCGCCACTGTCAATCGAAT	6282
Db	2936	GGCCCGGAGCGCTGTCTTACAAAGCCG	GAACACAGTCTGATGACACTTGGAGTGAAC	2995
QY	6283	CCGCGCGCTTACAGCGGTGTGCCAATG	ATGAATGAACACAGCTTTTGTGTCTTTCAGAAAGT	6342
Db	2996	GGCGCAGGCTTATGAGGATGCGAGATG	CAAGAACCGCGCTTTTGGGCTTTTCAGGAAGT	3055
QY	6343	GGCGGTACGATCAGCTTCCCGACAGCTG	CTTTTATCATCTTGGAAAGGTTTTCGAT	6402
Db	3056	GGCTATGTGCTATGATATCGCCCGGCT	TACACGCTTGTATCATCTGGAAGGCTTCGCAAT	3115
QY	6403	GGATGTGCGTGAAGAACGTTATGTGTG	TTTGGGGGACAGCGTGTGCTCTGTATCAAGT	6466
Db	3116	GGATGTATGCGAAGGCGCAATACAGCC	ACTGTGATATACGCTGTGCTTATTTGCTATCAAGT	3175
QY	6463	GGCGCGCGTGTGTGGGTTTAATGATG	CGCGCGTGTATGTGGCGGTACTGTATGAGCGGTA	6522
Db	3176	TGCAAGGGTGTTCGCGTTGATGATG	AGCGCAATCATGTGGCGGTGTGGGATTAACGCAAGCT	3235
QY	6523	CGATCAGCGCTGTGATTTGGGCTCTG	CGCTTTCAGCTTACGTTACCAATATGCAACGAGATATG	6582
Db	3236	GGAACCGCGCTGTGTGACTTTGGGCTG	GCAGTTTTCAGTTGACCAATATGTCTCGCGATATG	3295
QY	6583	TGAGAGACCGCGAGATGGCGGTGTGT	ACTGTGTGCAACAAAGCTGTGTGATGAGCGCGA	6644
Db	3296	GGAGCGATGCGATGTGGGGCGCGCTGT	ATATCTGTGCGGCAAGCTGTGTGAGCATGAAAGTCT	3355
QY	6643	GAGCGCGCCAGCTTGTGCGGATTCG	CAATTCGCAACAGCGCTGTGGCCCGCTGTGGACGCG	6700
Db	3356	GAAACAAAGAAATTAATGTGGGAC	CTGAAAAACGTCAGGCGCTGTAGCCGTATTCGCCCCGTG	3415
QY	6703	TCTGTGTGCGAGGCGGACCGGTACT	ATCATGCTCAGCGCGCAGCGGGGCTGTGCAAGATTTG	6766
Db	3416	TTTGGTGTGAGAAAGCAACTTA	CTTATTTTGTCTCCACAGCGCGCTGTGAGGGTTC	3475
QY	6763	GCTCCGTTTGGCGTGGGATTCG	CAACGCGCGCGGCTTTTACCGGGAAATTTGGGTAA	6822
Db	3476	CTGTGTGTGCGCTGTGGCAATTC	GTCTACGCGCAAGAGGTTTACCGGAAATTAATGGTGTCA	3535

[illegible]



FEATURES  
source  
1. .349107  
/organism="Photobacterium luminescens subsp. laumondii TT01"  
/mol\_type="genomic DNA"  
/strain="TT01"  
/db\_xref="taxon:243265"  
complement (80. .3225)  
/gene="tcccl1"  
/locus\_tag="plu4167"  
complement (80. .3211)  
/gene="tcccl1"  
/locus\_tag="plu4167"  
/function="unknown, probable insecticidal toxin"  
/note="Highly similar to insecticidal toxin complex protein Tccc of W14."  
/codon\_start=1  
/transl\_table=11  
/product="insecticidal toxin complex protein Tccc1"  
/protein\_id="CAE16539.1"  
/db\_xref="GI:36787442"  
/translation="MSTPDTALVTQTPVSVLIDNRGLSIDIGFHVIVIGSDITRYV  
RHQYDARYLASHSIDPRIDAKQADNSYKPFVQYDIAGALMTESYDAGRTYALND  
IEGRVMTWNAIVGRQTRKRGNTLPGKLSVSEQVPOQETSQVTERPTAMENTYAK  
EYNLSGLCTIRHYDTAGVTRIMSGIAGIILSQSHQWLAEGQANMSGDETVQRMIA  
SEVYTTQSTNAIGALLTQIDAKNIQGLAIDIAQOLKSGWLVKQGREQIVKSLSM  
SAAGHLEEHNGVITTEYSYEPETQRLMGITTRAESQSEARVLDLRYEYVGN  
VISIHNDAEATRFMRNOKVEBENRYVYDLSIQMSATGAREMANIGCOONOLPSAMPV  
PADNSTYNYQRTSYDRGNLVOIRHSSPATQNSYTTENTVSRSNRYTSLITTPD  
TVQGLFPAQGHQKTLIPQNLDMNMEKELRVIVPSENSSDEWYRSIDGRILK  
VSEQQTNSYTVQVRYVLPGLRLITGVADTQNTLVITVGEAGVQVRLVLEISRK  
PADIDNNQVRSYNNLSSQIELDSEGOILSQEEYYPYGGTAVMAAKQTEASFKYI  
RYSGERADATLGYYYGYQYQVGRMLSDPAGTVGLNVRVWNNPITLTDHGL  
ALPBNRNFTWFPASFLERKPDGMSAMRGOKIRAGIAGIILGIAAIACTAGA  
ALPITLGAIVAGAGIGALMGYNSLLEKGAIIARIVQSGTIVQSAAGAAASAA  
AAYARQGVGVASAGAVTGVASWLNADRGIGAGIAGSAGTITDMLTGISLTIL  
HEVAAGAAGAGMIITGQSTRAIGIAGITGYGSMIGFGLDVASVPAGLANVAG  
YVAGIGEMAIVRYGGFGLRGLRAVSPYAGIARQLVAFSIVARVFEIIFSVLGG  
LAGIGTGHVWRGRDWSVRGLSAGSGIDHLAGMIGNQIRGRVLTGTGIANADYG  
TSAIGARVIVSL"  
complement (3220. .3225)  
/gene="tcccl1"  
/locus\_tag="plu4167"  
complement (3345. .8051)  
/gene="tcccl1"  
/locus\_tag="plu4168"  
complement (3345. .8039)  
/gene="tcccl1"  
/locus\_tag="plu4168"  
/function="unknown, probable insecticidal toxin"  
/note="Highly similar to insecticidal toxin complex protein Tccb of W14 Putative transmembrane protein."  
/codon\_start=1  
/transl\_table=11  
/product="insecticidal toxin complex protein Tccb1"  
/protein\_id="CAE16540.1"  
/db\_xref="GI:36787443"  
/translation="MLSTMEKOLNESORDALVTGYMNFVASTLKGVDGQPVYEDLYE  
YLIDPRAVDEVERSVAAQALAIQOQWTRLVNGSEPRGQAMELSTANERDNDQQA  
IMAGAERYNAENYISPTTRQKSHFSELETTLNQRDLPPRYQAVAVIGYLEFSA  
VSNLYVLSGYINQKFPDQALYFLGRTTKRYRWRVQMDSKRQDPAGNPAVPPVNC  
NDWQBITLPLSGIDVLEHTVAPVFNDRLLVYAWWRPAPVQKADGKNIKTAHYSIK  
FGYKRYDDTWTAPNTTLMTOADESSTORSLLIGESKTLAKNMLATDTFSIDP  
TEETDSNPYGRMLGVFRANFTGRPHQNRPAVGYLICDSEFNHILRLPSKDELSPT  
YRDETNDQNCQPAVVDKRYVTKVTSIGKPEDEENGMCKVDLGLTGTGAYVYD  
QDGTTLHIQTTNGDPFNRRHFGYNDIVYDSKSGYGTWGSNEFTYLDHGNYYTH  
NAALNYVAGYGGGSPVNGTWALRQVNEGALSPLLDHHTTIVQSGYLAWEGEYTA  
SYVLPIDGTMLDMSDKINFALISLNTKESVTSPPDWPTLTITKPSKIXADRFTQY  
INAEIADGRNLFKHSQTFOLASGATPYTTLSEADFSIDPDKXVLCIACLVNVMH  
YDRSGRKGASVSKWENVYVELDQSKAPAIPLISRYDSKRGLOVYDDEPTSSIP  
AKRILNTFVATLIEKANIGLDSLDTLQADPSLEADIVTQDSEPMDDGSGNGIYF  
WELFPHLPEVATRFANEQSPKQSLHITFDPMKKPHNAPYANVRPLVEYHSD  
LSRLGDPIDPDTQAVAHVPIYQCAVSIYAVSNLIAQGDMMYRQJLTDGALQAVVYN  
LAELLPBRPDVSLSTWTPTQTLDTLAAQKTLARDFEHOIANSNTVLPAIPGNINISH

gene  
RBS  
CDS  
1. .349107  
/organism="Photobacterium luminescens subsp. laumondii TT01"  
/mol\_type="genomic DNA"  
/strain="TT01"  
/db\_xref="taxon:243265"  
complement (80. .3225)  
/gene="tcccl1"  
/locus\_tag="plu4167"  
complement (80. .3211)  
/gene="tcccl1"  
/locus\_tag="plu4167"  
/function="unknown, probable insecticidal toxin"  
/note="Highly similar to insecticidal toxin complex protein Tccc of W14."  
/codon\_start=1  
/transl\_table=11  
/product="insecticidal toxin complex protein Tccc1"  
/protein\_id="CAE16539.1"  
/db\_xref="GI:36787442"  
/translation="MSTPDTALVTQTPVSVLIDNRGLSIDIGFHVIVIGSDITRYV  
RHQYDARYLASHSIDPRIDAKQADNSYKPFVQYDIAGALMTESYDAGRTYALND  
IEGRVMTWNAIVGRQTRKRGNTLPGKLSVSEQVPOQETSQVTERPTAMENTYAK  
EYNLSGLCTIRHYDTAGVTRIMSGIAGIILSQSHQWLAEGQANMSGDETVQRMIA  
SEVYTTQSTNAIGALLTQIDAKNIQGLAIDIAQOLKSGWLVKQGREQIVKSLSM  
SAAGHLEEHNGVITTEYSYEPETQRLMGITTRAESQSEARVLDLRYEYVGN  
VISIHNDAEATRFMRNOKVEBENRYVYDLSIQMSATGAREMANIGCOONOLPSAMPV  
PADNSTYNYQRTSYDRGNLVOIRHSSPATQNSYTTENTVSRSNRYTSLITTPD  
TVQGLFPAQGHQKTLIPQNLDMNMEKELRVIVPSENSSDEWYRSIDGRILK  
VSEQQTNSYTVQVRYVLPGLRLITGVADTQNTLVITVGEAGVQVRLVLEISRK  
PADIDNNQVRSYNNLSSQIELDSEGOILSQEEYYPYGGTAVMAAKQTEASFKYI  
RYSGERADATLGYYYGYQYQVGRMLSDPAGTVGLNVRVWNNPITLTDHGL  
ALPBNRNFTWFPASFLERKPDGMSAMRGOKIRAGIAGIILGIAAIACTAGA  
ALPITLGAIVAGAGIGALMGYNSLLEKGAIIARIVQSGTIVQSAAGAAASAA  
AAYARQGVGVASAGAVTGVASWLNADRGIGAGIAGSAGTITDMLTGISLTIL  
HEVAAGAAGAGMIITGQSTRAIGIAGITGYGSMIGFGLDVASVPAGLANVAG  
YVAGIGEMAIVRYGGFGLRGLRAVSPYAGIARQLVAFSIVARVFEIIFSVLGG  
LAGIGTGHVWRGRDWSVRGLSAGSGIDHLAGMIGNQIRGRVLTGTGIANADYG  
TSAIGARVIVSL"  
complement (3220. .3225)  
/gene="tcccl1"  
/locus\_tag="plu4167"  
complement (3345. .8051)  
/gene="tcccl1"  
/locus\_tag="plu4168"  
complement (3345. .8039)  
/gene="tcccl1"  
/locus\_tag="plu4168"  
/function="unknown, probable insecticidal toxin"  
/note="Highly similar to insecticidal toxin complex protein Tccb of W14 Putative transmembrane protein."  
/codon\_start=1  
/transl\_table=11  
/product="insecticidal toxin complex protein Tccb1"  
/protein\_id="CAE16540.1"  
/db\_xref="GI:36787443"  
/translation="MLSTMEKOLNESORDALVTGYMNFVASTLKGVDGQPVYEDLYE  
YLIDPRAVDEVERSVAAQALAIQOQWTRLVNGSEPRGQAMELSTANERDNDQQA  
IMAGAERYNAENYISPTTRQKSHFSELETTLNQRDLPPRYQAVAVIGYLEFSA  
VSNLYVLSGYINQKFPDQALYFLGRTTKRYRWRVQMDSKRQDPAGNPAVPPVNC  
NDWQBITLPLSGIDVLEHTVAPVFNDRLLVYAWWRPAPVQKADGKNIKTAHYSIK  
FGYKRYDDTWTAPNTTLMTOADESSTORSLLIGESKTLAKNMLATDTFSIDP  
TEETDSNPYGRMLGVFRANFTGRPHQNRPAVGYLICDSEFNHILRLPSKDELSPT  
YRDETNDQNCQPAVVDKRYVTKVTSIGKPEDEENGMCKVDLGLTGTGAYVYD  
QDGTTLHIQTTNGDPFNRRHFGYNDIVYDSKSGYGTWGSNEFTYLDHGNYYTH  
NAALNYVAGYGGGSPVNGTWALRQVNEGALSPLLDHHTTIVQSGYLAWEGEYTA  
SYVLPIDGTMLDMSDKINFALISLNTKESVTSPPDWPTLTITKPSKIXADRFTQY  
INAEIADGRNLFKHSQTFOLASGATPYTTLSEADFSIDPDKXVLCIACLVNVMH  
YDRSGRKGASVSKWENVYVELDQSKAPAIPLISRYDSKRGLOVYDDEPTSSIP  
AKRILNTFVATLIEKANIGLDSLDTLQADPSLEADIVTQDSEPMDDGSGNGIYF  
WELFPHLPEVATRFANEQSPKQSLHITFDPMKKPHNAPYANVRPLVEYHSD  
LSRLGDPIDPDTQAVAHVPIYQCAVSIYAVSNLIAQGDMMYRQJLTDGALQAVVYN  
LAELLPBRPDVSLSTWTPTQTLDTLAAQKTLARDFEHOIANSNTVLPAIPGNINISH



```
AKGMLAPKRDGKILPAVELFLIDOKRPLYOFPNELPSLKEKPLMTWVWHBOFGKRE
KALPRKRLAKDYVWTELYTLDFRIPTSMQVYDILBOVKSAVFLSTSEVLM"
gene
  complement (12597..13081)
  /locus_tag="Plu4172"
  /locus_tag="Plu4172"
  /locus_tag="Plu4172"
  /locus_tag="Plu4172"
  /function="Unknown"
  /note="unnamed protein product; Similar to pyocin S3
  immunity protein and to Photobadus luminescens proteins"
  /codon_start=1
  /transl_table=11
  /protein_id="CAE16544.1"
  /db_xref="GI:36787447"
RBS
  /translation="MIDVIEIINSLEKTDMLVENKILPEKPEFPEGSRSVPE
  TGLIVDILSVISVQFLINVDVTGYSDMDPAPLHKRNKAIVRAPWGEPPYS
  GSEPRIPVIGLVGFPDYTRKLDQYPTNEIRFLYOTDLRVDLALFEOSVND"
  /complement (13076..13081)
  /locus_tag="Plu4172"
  /locus_tag="Plu4172"
  /complement (13174..13652)
  /locus_tag="Plu4173"
  /complement (13174..13635)
  /locus_tag="Plu4173"
  /function="Unknown"
  /note="unnamed protein product; Some similarities with
  pyocin S3 immunity protein and similar to Photobadus
  luminescens proteins"
  /codon_start=1
  /transl_table=11
  /protein_id="CAE16545.1"
Query Match      15.5%; Score 1345.8; DB 1; Length 349107;
Best Local Similarity 61.5%; Pred. No. 1.3e-242;
Matches 2191; Conservative 0; Mismatches 1367; Indels 5; Gaps 2;

QY 3421 GATGCAATTCGTGCGTGGCGCGCTGCTAAAGCGCGTGAATGCGCTTAAAGCGCAG 3480
DB 202949 GATCTGATTCGTGCGGAGAGATTTGGCCAGCATGATCCGATTCGCTTCCAGCAG 203008
QY 3481 CTGACGCGCGCATTTAAGTTTGTCTAATGAGAGTCAAGCGCGCGCGCGCATCAT 3540
DB 203009 TGTAAACCATCTGCGAGAGTGTCTAATGAAAAACAGAAACAAATAGAGGCAATCAC 203068
QY 3541 ACCTGCTGCTTCATGCGGAGACGTCAAGGAGCGGAGTTTGGCTGCTGAGCGCGCTG 3600
DB 203069 ACCTGCTGCTTCATGCAACATGATCTTACGAGCGGAGCATGATGATAGACGCGCTG 203128
QY 3601 CTTTGGCGCGCTGCGCGCGCTTATCAGGTACGCTTCCCGACCTGCGTCCGCGCATGAT 3660
DB 203129 ATTACCTATGCTGCTGAGGTTCAGAGCTCATTTTCCAGCATTTTCAACGCAATTCCCA 203188
QY 3661 GGTGAATATTGCTGATTCGCTGAGAGATTTTGGCGCGCATTTACAGAGGTGCTCGAT 3720
DB 203189 CATTCAATATTCAAGTATCATATCCCAACATTTGCAAGCATATCCATGATATTGGGC 203248
QY 3721 GCGCGCTAGACACCGGAGCGCGCTGAGAGGTCTACCCACCGGAGTCAAGTGGC 3780
DB 203249 GAACTGATACAGACTGCTTATTATGTAACAGAGCTGACTCCACAGAAAGTTTACTTACAG 203308
QY 3781 GATGCGGAGATTTTACAGCGCAGCGGATTTGACCGGCGCGGCTGACCGCAGCAGCG 3840
DB 203309 GACGGCTGCTCTTAAGTGTGCGCGCATTTGATGATGAGGAGGTGGCCACAGAGCA 203368
QY 3841 CATTCGAGCTGCGCTATCAGGATTTGTGCTCAGAGATGGCAACTGCGCGCGCGCAT 3900
DB 203369 TTTATAGGAGTGGACCCAGCATTTTGTTCAGAGATGGAGCTGGAAGAGTGGCAC 203428
QY 3901 GCGCTGAGAGAGCAATATGATGAGACGACAGGCTCATAGACAGCGGTTTATCGCTT 3960
DB 203429 TCTTTAACCAACCGATTTTAAATGATATAGTGTGAGCAAGATACAGGTTTATCATTTT 203488
QY 3961 GTTTAACACCTGCGCTGAGTGCAGCGTTTACTGATTAAGATACCACTCATCAATCAAC 4020
DB 203489 ATCTAAGTCTGCTGCTCTCTCAACTGCTGCTATAGAGACACTCATTAAGTTGAT 203548
```

QY	5099	GGATCATCTGGCGCAGGCGCTTTTTCCTTTACATCGTTGCTGGCGGGATACCTTTTGC	5158
Db	204629	TGAGAGCTCGCGCCAGCGCAATTTTCATTTCATCTCGCTATTGATAGAGGATACCCCTTTGC	204688
QY	5159	AACGCTTCGATCTATACCTTAATTCACCGCACCTGAGCGCGAATGGGCGTGGTTTCC	5218
Db	204689	TGCATCTTCTATCTATCTACTTTTATCCAGCCACTGGAAACGAAATGGGGAGTCTGGTTTCC	204748
QY	5219	GGCGGCGGTACCGGTGCGCTGGTTGATATGGCATATGCGCGGCTGTTTCGCAATTTGGGCGG	5278
Db	204749	GCGAGAGAGAACACAGCGCTCTGTGGTTGAGGCATATGGTAATTGTTTACTGACATTTGTGTGG	204808
QY	5279	TGAACTGCTGCTCAACGCCGGAATGACGCCAGCTGGAGAACGAGGGGTAACCGCATACAGCGG	5338
Db	204809	AGAAATTAAGCTTAATATGCAAAAGTATAAAGCTTTCACACCCCATAGGCAATATGGGTGCACAGG	204868
QY	5339	TGTCAGCTGAAGGATATGGGCGCGGTTTTGCCCGCGCGCGCTTGCGCTCAATCTGACGT	5398
Db	204869	TGTCCAATTAGCTGATGAGCAGAAATATGACATATGACATTTGTGGCCTCAAAATCCGATGT	204928
QY	5399	GGTGCATACCTACGATATCGCTGTATTAAGCAGCATCTCTGCGGCGGTAAACGCGGCGCAAC	5458
Db	204929	TATTCAATACCTATTAAGCACTGTATAGTACAGCATCCGATCGGAATATAGCCGCGCTGAAAC	204988
QY	5459	GCTGAAGGGCAAGCGGATGAGCAATCGCTGTTTGTACTCTATTTTGGTCTTATATATATGC	5518
Db	204989	ACTGGATGGTAAACGAATATAGCAATCACTGTTCGGCTTTATTTTGGTCTGGAACATATCA	205048
QY	5519	CCACCGCAGCTGGCGGCAACACACGCGTGTGCTTGGTCCGCGCTATCGTAATTGATATCGA	5578
Db	205049	CCATATCGCAATTAGCCCAATCAATACAGTTTGTCTTTCCACGCTATTAAGGAATCAATCGA	205108
QY	5579	TGAGATCTTCAATTAACAGCGACGACTGGCGGAAGATTTCTCGCTGTATCTGCATATGCGCTTG	5638
Db	205109	AGATATTTTATTAATCAATGACCGCACTGTCAAGAGATTTTTCATCTATCTCATATGCCCTTC	205168
QY	5639	CTCCAGCGAATCCGTGCTGTGGCAACCGGCGGGGCTTGGCGCAATTTTAACTGTCTGGCGCGGT	5698
Db	205169	TGTTATCTAACCCTATCCCTTGTCTCCCGAGAAATGGCGAAGTATCTAAGTTCCTGGCCCTGT	205228
QY	5699	GCGCGATCTCGGATACCGCGCGCAATTGACTGGGCAACGAGAAAGGCGCGCTGGCGCATGCG	5758
Db	205229	ACCAATTTTGGGGAAGCTAATTTTAACTGGGATATAGAGGGGCCACGCTTACGTACACG	205288
QY	5759	CATTTTCTTATCTGAGAGACATATAATGCCGGGTCTGCGACAGCAGTTATGTACACAC	5818
Db	205289	TATTTTCCGGTATCTGAAAAGTACATATATGCCCGGTCTGCTCAAGCAGTTATGTATCGTTCA	205348
QY	5819	CCGTAATGTTTACGCCGTTTGAATTTTGGCAACAGCTGACACGCGCATACAGGGCTCTACGCTT	5878
Db	205349	TGTAATTTTATCCCAATTTGATTTTCTGTACCAATCAATATGCCCATCTCGGTTCGGCCTT	205408
QY	5879	TTTCGTCGAACCATTTTGAAGCAAAAGGCGCTGGTTCGGGCGCATTAACGCGATGCGCA	5938
Db	205409	CTTCTTTGAACTGCTGTCTGACACAAAGCGCTTGTGTTCCGACACACAAATGTGGACAACTCG	205468
QY	5939	CATTACTAACCTTTATCTGTGGGGGCTGGCAACGATCCCGGTTGCCGATGTGCCAGCGT	5998
Db	205469	CATTATATATCTGTATCTCTGTGGGCGAGGCAACGATCTCGAGGAGATATTCGCGGCGT	205528
QY	5999	GATCGGCTCCGCGAAAGCAGCCGCCAGCTGATGTGAGAGATCTGAACGGATGAACCA	6058
Db	205529	GATTGGAATCAGCAAAAGCCACCGCCACATTTGATGTTAGAGATATGCTTAAATGA---AT	205588
QY	6059	CCGCGGCTGATGAGAGGTACACGCAACATATGGGCGCAAGGCTCCAAAGTTTCCGACAC	6118
Db	205586	CTTGCGCTACTTAAACAGCTTACTCAGATTAATGAGACAAAGGCTCAAAAGTTTTCACAGT	205645
QY	6119	GCTACCGGCTATTTGATCTTCAACGCGCGCAGTACGCTGATGCTGTACGCGCTGATG	6178
Db	205646	GTCACCCCACTTTTGGATACAGACACACACACATATGATGCTGTACGCGCTGTGTC	205705
QY	6179	CGTACTGTGACGATGTGTATGATGTGTACAGCTGTGGCGAAAGCGGCAACGACGACGCG	6238

Db	205706	CGTATTGTGATGATTAATATAGACGGGCAAGAGCTGGGAGACAATAATAGCAGTGTGAT	2057655
Qy	6239	GTGGCGGATGACACAGGCGCGGATGCGCACTGCAAAATCGAAACCCGCGGCTTAACAGC	6298
Db	205766	AAGTATAGCCGCCGCGAGAAACTTCAATGCTGCAATACCTGACAAGACAGGCTTTCGAC	2058255
Qy	6239	GGTGCCCACTAGATGAACCAAGCTTTCGTGCTTTCAGGAAGTGGCGGTGACATCAG	6358
Db	205826	GGTTTTCGGAATGACCGAACTGCTTTTTCACGCTTTCAGACAGTGAAGCTCTGAGTAATGAA	2058858
Qy	6359	CTTCCCGACAGCTGCGCTTTTGATCATCTGGAAGGTTTTCGATGGATCCGCGTGAAGAA	6418
Db	205886	ATCCTCAGCAACAGGCTTTCGAAACATCTGGAAGGATTCGCAATGGATTTACTCTGTAG	2059455
Qy	6419	CGTTATGCGTGTTCGGGGAGACCGCTGCGTTACTGCTATACGTTGCGCGGCTGTGGGG	6478
Db	205946	CCTTATACGAAGCTTGTAGATGATATCTGTAGATATTTCTACATGCTCGCGGAGATGATTTGGG	2060050
Qy	6479	TTAATGATGCGCGCGGTATGAGGCTAGCTGATGAGCGGTACTGATACCCCTGTGAT	6538
Db	206006	CTGATGATGCGAAGGTTATGGGGTTTCGCAAGGAGTGTCTGATATGGGCTTCGCGAT	2060655
Qy	6539	TTGGGTGCGGGCTTTCAGCTTACCAATATCGACCGGATATCGTTGAGGACCGCGAAT	6598
Db	206066	CTAGGAATTCCTTTTCAATTATCTAATACATTTGCTCGGATATTTATGAGATTCGAAGGCC	2061255
Qy	6599	GGCGGTTGCTATCTGCGCAAAAGCTGCGTGAATGAGGCGGACTGAGCCCGCCAGCTT	6658
Db	206126	GGGGGTTGCTATCTTCGCTGAAATGGCTTTCACAAAGAGGCTAATGCGACACACTT	2061855
Qy	6659	GCCGATCCGCAACATCGCGCAGCGCTGGCCCGCTGCGCAGCGCGTCTGTGGCGCAGGCC	6718
Db	206186	ATTATATACGAAACCGTCTCGCTGACTCGGGTGAATCCCGATTGATTTGGAAGGG	2062455
Qy	6719	GAGCGGTATCTATCAGTCAAGCGCGACGCGGCTGCCAGATTCGCGCTTCGCGGCTGG	6778
Db	206246	GAATTTATTTATCTCCGCTCTGACAGGGCTTTGTTGGTTTACCTTTGCGTTTCAGCTGG	2063050
Qy	6779	GCGATTCGCCAACCGCGCGCGGCTTTTACCGGAAATTTGGCGTAAAGTGCAGCATGCGGT	6838
Db	206306	GTCAATCGCTTCTGCTCATGATATCTATGTAAGTCCGATTTAAAGTGCACAACAGAGGA	2063655
Qy	6839	GCCCGGCGATGGATACCGCGCACGCGCAACAGTAAAGCGGAAAGCTGGCGCTCTCTGGTG	6898
Db	206366	GTACGCGGCTCTGGGAATGCAGACAAAGAACTAATCGTGGAGAAAGATTACCTTATTTAATG	2064255
Qy	6899	AAAGGTGCGGCGGTGCGGCTTACTTTCGCGGCTTGTGCAATCCGAGGCGGCTCTCGCGGT	6958
Db	206426	GCTGGCGGAATGAAGCGCTTATATCTGGGTTGGGATPAACAACGCTTCGCAATCCAAA	2064855
Qy	6959	CTGTGCGAGCGTCCGCGTTGACA	6981
Db	206486	TTGTGCGACGACCTCATTTAAA	206508
RESULT 9			
ERMCRTRIB	2415 bp	DNA	linear - BCT 11-APR-2001
LOCUS			
DEFINITION	Pantoea agglomerans phytoene dehydrogenase (crtI) and prephytoene		
	pyrophosphate synthase (crtB) genes, complete cds.		
ACCESSION	M38423		
VERSION	M38423.1	GI:148401	
KEYWORDS			
SOURCE	Pantoea agglomerans		
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
	Enterobacteriaceae; Pantoea.		
REFERENCE	1 (bases 1 to 2415)		
AUTHORS	Armstrong, G.A., Alberti, M. and Hearst, J.E.		
TITLE	Conserved enzymes mediate the early reactions of carotenoid		
	biosynthesis in nonphotosynthetic and photosynthetic prokaryotes		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 87, 9975-9979 (1992)		

MEDLINE 91086634  
COMMENT Draft entry and computer-readable sequence for [Proc. Natl. Acad. Sci. U.S.A. 87, 9975-9979 (1990)] kindly submitted by M. Alberti, 11-SEP-1990.

FEATURES  
source location/Qualifiers

1. .2415  
/organism="Pantoea agglomerans"  
/mol\_type="genomic DNA"  
/strain="rho10"  
/db\_xref="taxon:549"  
11. .1489  
/gene="crtI"  
11. .1489  
/gene="crtI"  
/note="putative"  
/codon\_start=1  
/transl\_table=11  
/product="phytoene dehydrogenase"  
/protein\_id="AA24820.1"  
/db\_xref="GI:148402"

CDs  
gene  
1486. .2415  
/gene="crtB"  
1486. .2415  
/gene="crtB"  
/note="putative"  
/codon\_start=1  
/transl\_table=11  
/product="prephytoene pyrophosphate synthase"  
/protein\_id="AA24821.1"  
/db\_xref="GI:148403"  
/translation="MSQPLLDHATQTMANGSKSPATPAALFDPATRRSVLMYTCR  
HCDVDLDQTHGPAEAAEBAEATQRLARLTLLAFBAEMODPAFAFOVALTH  
GITPRALDHLDFAMDAEATQRYVTFEDTLRCYHVAQVGLMARVMGRVDRVDR  
ACDLGARFQLTNTARDIIDAIDRCYLPBEMQDAGTPENYAAENRALAVAR  
LIDAEPTYSQAQGLHDLPPRCAMATIKRSYREIGIKVKAAGSAMDRQHTSKG  
EKIAMMAAGQVIRAKTYRTPRPAGLMQRPV"

## ORIGIN

Query Match 14.6% Score 1265.8; DB 1; Length 2415;  
Best Local Similarity 70.6% Pred. No. 2e-227; Indels 3; Gaps 1;  
Matches 1703; Conservative 0; Mismatches 707;

QY 4565 AGAATATGAAAGCAGCCTATGATGATGCGAGCGTTCGTCGCGCTGCGAT 4624  
DB 4 ATAAAGATGAAAAAACCCTTGATGATGCGAGCGCTTGCGCTGCGCTGCGAT 63  
QY 4625 TCGTTCGCAAGCGCGCGCTGCGCTGCGCTGCGAAACAGCGCATTAAGCTGCGG 4684  
DB 64 TCGCGTCAAGCGCGCGAGGATCCCAACCGTATGCTGAGACAGCGGGAACCGCGCG 123  
QY 4685 GCGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4744  
DB 124 TCGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183  
QY 4745 CGATCCAGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4804  
DB 184 CGATCCAGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243  
QY 4805 CGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4864  
DB 244 CAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303  
QY 4865 CTACGACAAATCAGCGCGAGCTGAGCGAGCGATGCGACCTTTTATCCAGATGT 4924

DB 304 CTATGCTAAAGCAGCGCGAGCTTGAAGCGCGATTAACCACTTCAACCCCGCGAGT 363  
QY 4925 CGCGGCTTACCGGAGCTTCTGCGCTTATCAAGATGTTTTCGAGGCGTATCTGA 4984  
DB 364 CAGAGGCTACCGGAGCTTCTGCGCTTATCAAGATGTTTTCGAGGCGTATCTGA 423  
QY 4985 ACTGAGGACCGGAGCTTCTGCGCTTATCAAGATGTTTTCGAGGCGTATCTGA 5044  
DB 424 CTGCGGAGGCTTCTGCGCTTATCAAGATGTTTTCGAGGCGTATCTGA 483  
QY 5045 TCGGCTGCAAGCGCTGCGAGTGTCTACAGCAAGTGTGCGAATTTATTCAGATCA 5104  
DB 484 TAGCTTCAAGCGCTGCGAGGCTTACAGCAAGTGTGCGAATTTATTCAGATCA 543  
QY 5105 TCTGCGCAGGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5164  
DB 544 TCTGCGCAGGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 603  
QY 5165 TTGATCTTATCTTATTCAGGCACTGAGCGGAGTGGGCGGTGTTTCGCGCGG 5224  
DB 604 GTCCATCTACACCTTATTCAGGCACTGAGCGGAGTGGGCGGTGTTTCGCGCGG 663  
QY 5225 CGGTACCGGTGCGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 5284  
DB 664 CGGCAACCGGCGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 723  
QY 5285 GCTGCTCAACCGGAGTCAAGCAAGTGTGATGATGATGATGATGATGATGATGAT 5344  
DB 724 CGAATCTCAACCGGAGTCAAGCAAGTGTGATGATGATGATGATGATGATGATGAT 783  
QY 5345 GCTGAAGATGAGCGCGCTTTGCGCGCGCGCGCGCTTGAAGTGTGATGATGATGAT 5404  
DB 784 GCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843  
QY 5405 TACCTAGATGCGCTTGAAGCGGATGATGATGATGATGATGATGATGATGATGATGAT 5464  
DB 844 CACTTATTAAGCTGCTGCGGCAACCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 903  
QY 5465 GCGCAAGCGGATGAGCACTGCTGTTTGTACTCTATTTTGTGCTTATCATGCCAACC 5524  
DB 904 GCGCAAGCGGATGAGCACTGCTGTTTGTACTCTATTTTGTGCTTATCATGCCAACC 963  
QY 5525 GCGAGTGGCGCAACAGCGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 5584  
DB 964 CAGCTGGCGCAACAGCGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1023  
QY 5585 CTCAATGAGCAGCAGCTGCGGAGATTTCTGCTGATCTGATGCTGCTGCTGCTGCTGCT 5644  
DB 1024 CTCAATGAGCAGCAGCTGCGGAGATTTCTGCTGATCTGATGCTGCTGCTGCTGCTGCT 1083  
QY 5645 CGATCCGCTGCTGCGGAGCAGCGGCTGCGGAGATTTTACGTCGCGCGCGGCTGCGCA 5704  
DB 1084 CGATCCGCTGCTGCGGAGCAGCGGCTGCGGAGATTTTACGTCGCGCGCGGCTGCGCA 1143  
QY 5705 TCTGCGTACCGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5764  
DB 1144 TCTGCGTACCGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203  
QY 5765 TCGTATTCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5824  
DB 1204 TCGTATTCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263  
QY 5825 GTTTAAGCGCTTGTATTTTCGAGCAAGCTGCAAGCGGATCAGGCGCTTTCGCT 5884  
DB 1264 GTTTAAGCGCTTGTATTTTCGAGCAAGCTGCAAGCGGATCAGGCGCTTTCGCT 1323  
QY 5885 CGAATCCATTTTGAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5944  
DB 1324 CGAATCCATTTTGAAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1383  
QY 5945 TAACTTTATCTGATGAGGCGCTGCAAGCTGCAAGCGGATCAGGCGCTTTCGCT 6004  
DB 1384 TAACTTTATCTGATGAGGCGCTGCAAGCTGCAAGCGGATCAGGCGCTTTCGCT 1443



## ORIGIN

NDIPERYPLQLEGRMDVEGRYDILDTLYCYHVAGVGVMMALINGARDEVLD  
RASLDLGFQUTINARDVEMDAIGRIYLAALFVAGVBEAIAAEHHAAYAVYA  
RLIDVNEPYDOLALIGMAALPRBAALFVAGVBYRAIIGIEVNRKGRKANDTVSTST-  
AOKRAGFLAGLQALATRFRCAPPFPAHMLTRPA"

Query Match 13.3%; Score 1156.4; DB 1; Length 3801;  
Best Local Similarity 58.9%; Pred. No. 7e-207;  
Matches 2123; Conservative 0; Mismatches 1436; Indels 45; Gaps 6;

QY 3421 GATGTCATTCGTGCGTGCCTGCGCTGCTAAACGGGCTGATGTGGCTGCTTAAACGCCAG 3480  
DB 173 GACATCTCTTCGCGCGCGGCGGACTGCCAATTGCTTAATGCGGCGCGGCTCGCGACC 232  
QY 3481 CTGAGCGCCGCACTTAAGTTTGGTCTACTGAGAGTGAAGCGGCGCGCGCGCAATCAT 3540  
DB 233 CAGGCTCCCGGCTTCAATGCTGCTGCTGAAACCGCGCGAGACGCGGCGGCAATCAC 292  
QY 3541 ACCGTGCTTCCATTCGCAAGACGTGACGCGAGCGGATTCGCTGGCTGAGCGGCTG 3600  
DB 293 AGCTGTCTGCGCAAGACGAGATCTAATGCGCGAGCGCGCTTCTGCTCCCTTC 352  
QY 3601 CTTTGGCGCGCTGCGCGCGGTTTACAGTACGCTTCCCACTGCGCTGCGCAAGTGAAT 3660  
DB 353 CAGAGCTATCTGCGCGCGGCGGAGGTGATTTTCCGCTTTTTCGCGAGCGTCAA 412  
QY 3661 GGTGAATATTCGTGATTTGCTCGGAGATTTTCCCGCGCACTTAACGAGTCTCGGT 3720  
DB 413 GGGGCGCTACGCAACATTTCTTCGAGCGCATGCGCGAGTGAATGAGCGCTGTGCG 472  
QY 3721 GCGCGCTACGCAACCGGAGCGCGGCTGAGGAGGCTCACCCAGCGGAGTCAAGCTGGCG 3780  
DB 473 GCGGCTACCGGCAACATTCGCGCGGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 532  
QY 3781 GATGCGCGGATGTTACAGCGCGAGCGGATGTTACAGCGCGCGGCTGACCGCAACCG 3840  
DB 533 GGGGCGGAGGCAATGAGCGCGCGCGCGGTGAGAGCGGCGCGCGCTGCGCGCTGCGCG 592  
QY 3841 CATCTGAGCTCGGCTATCAGGCAATTTGTGCTGATGAGAGTGGCACTGGCGCGCGCAT 3900  
DB 593 CATCTGAGCTCGGCTATCAGGCAATTTGTGCTGATGAGAGTGGCACTGGCGCGCGCAT 652  
QY 3901 GGCCTGAGGAGCAATATGATGAGAGCGAGCGGTGATGAGAGCGGATTTTCGCTT 3960  
DB 653 GGCCTACCGCGCGCATATGATGAGAGCGGAGTGAAGAGCTGGCGGTTACCGCTTC 712  
QY 3961 GTTTACACCTGCGCTCAGTGGCGCGGTTTACTGATTAAGATACCACTACATCAAC 4020  
DB 713 GTCTATGTGCTGCGCTGAGAGATGACACGCTGTGATGAGAGACCTATTAATGCGGAC 772  
QY 4021 CATGCCAGCTGATGAGCGCAAGGCGCGCGCTCACTTAACGATTAATGCGGACGCGC 4080  
DB 773 GGCCTGAGTCTTCCCGCGAGTGCCTTGGCGGCGGCAATTCGCGCTATGCGCGCGGAG 832  
QY 4081 GGCCTGAGTCTTGGCGCGGCTGCTGCGGAGAGACGCGCTGCGCGGATTAACGCTACG 4140  
DB 833 GGCCTGAGTCTTGGCGCGGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 892  
QY 4141 GGCAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4197  
DB 893 GGCAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952  
QY 4198 GCGGAGCTGTTTACGCGCAACCGGTTTACTGCTGCGCGCGCGGCTGCGGAG 4257  
DB 953 GCGGAGCTGTTTACGCGCAACCGGTTTACTGCTGCGCGCGGCTGCGGAG 1012  
QY 4258 AAGATTCGAG 4317  
DB 1013 TCGGTGAGCGCGCTTTCGCGCGCGCGCTTTCGCGCGAGTGCAGGCGCAT 1072  
QY 4318 GCGGCTGAG 4377  
DB 1073 GCGGCGCGCGGAG 1132

QY 4378 GCGGAGAGGCTGAACAGCGGCTGCGCGGCTAATGAGCGCTTTTACGCGGTTGAGCGCGGA 4437  
DB 1133 GCTGCGCATCCGAGCGGCGGCTATGCTATCTCAAGGCTTCTCAAGGCGCTGTCGGAAGAC 1192  
QY 4438 TTGATTAAGCGCTTTTACGCGGCGCACTGCGCTGAGCGATTAAGACGATTCGTGCG 4497  
DB 1193 CTGATGCGCGGTTTCTAAGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1252  
QY 4498 GCGCAACCGCGGCTCTCTGCGGAGAGCGCTGCGGCTGATGATGAGCTCTCGCTTA 4557  
DB 1253 GCGCGCGCAACCGGATGAGGTTTCAAGGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 1312  
QY 4558 CCAAGGAGAAATTAAG-----AAAGCACTATATGATTTGGCG 4596  
DB 1313 CCGGCTTCCCATGATGCTGATGCAACAAGACATCCCGCGCGCGCTGCTGCTGCTGCTG 1372  
QY 4597 CAGGCTTGGCTGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4656  
DB 1373 CCGGCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1432  
QY 4657 TGTGGAACAGCGCGATTAAGCTGCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4716  
DB 1433 TGTGGAACAGCGCGATTAAGCTGCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1492  
QY 4717 CTTTGAATGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4776  
DB 1493 CTTTGAATGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1552  
QY 4777 TGGGAGGCAAGCAATCAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 4836  
DB 1553 CCGCGGCGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1612  
QY 4837 TGTGCGGAGAGAGCGGAGGAGCTGAGCTAGCAACAATGAGCGGAGCTGAGAGAGC 4896  
DB 1613 TGTGCGGAGAGAGCGGAGGAGCTGAGCTAGCAACAATGAGCGGAGCTGAGAGAGC 1672  
QY 4897 AGATTCGCACTTTTAATCCCAAGAGTGTGCGCGGCTTACCGCGGCTTCTGCGCTATTCAC 4956  
DB 1673 AGATTCGCACTTTTAATCCCAAGAGTGTGCGCGGCTTACCGCGGCTTCTGCGCTATTCAC 1732  
QY 4957 AGATTCGCTTCTGAGAGGCTATCTGAAACTGAGGAGCGGATCTTCTGCTGCTGCTGCTG 5016  
DB 1733 AGGCGCTTCTGAGAGGCTATCTGAAACTGAGGAGCGGATCTTCTGCTGCTGCTGCTGCTG 1792  
QY 5017 ACAATGCTGCTGCGGCGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5076  
DB 1793 GCAATGATGAAGCGCGCGCGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1852  
QY 5077 TGTGCGGAAATTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5136  
DB 1853 TGTGCGGAAATTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1912  
QY 5137 TGTGCGGCGGATTCCTTTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 5196  
DB 1913 TGTGCGGCGGATTCCTTTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1972  
QY 5197 GCGAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5256  
DB 1973 GCAATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2032  
QY 5257 GCGCTGTTTGGGATTTTGGGCGGTTGAACTGCTGCTCAACCGGAGTGAAGCTGAGAG 5316  
DB 2033 AGCTGTTTACCGGACTTGGGCGGAGATCGGCTTCAACCGGAGTGAAGCTGAGAG 2092  
QY 5317 CCGAGGATTAACGATCAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 5376  
DB 2093 TGAAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2152  
QY 5377 CCGTTGCGTCAAAATGCTGAGAGTGTGCTATACCTAGAGCTGCTGTTAAGCGAGCATCTG 5436  
DB 2153 TCGTCCGCTCAATGAGAGTGTGCTCAACCTATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2212



QY 5437 GCGGCGGTAACGCGCGCAACGCTGAAGCGCAAGCGGATGAGCACTCGCTGTTGATC 5496  
 Db 2213 GCGGTCGTGGAAAGCGCGCGCTGGCCAAAGACGCCCACTCATATGCTGTTGCTCA 2272  
 QY 5497 TCTATTTTGGCTTTAATTCATGCTCCACCGCAGCTGGGCGACACAGCGTGTGTTGGTC 5556  
 Db 2273 CTTATTTTCGGGGCGCGCGGACCTGGGACCATCTCCAGACACACAGGTGTGTTGGTC 2332  
 QY 5557 GCGGCTATTCGTGAATTGATGATGATGATCTTCAATAGCAGCAGCTGGCGGAATTTCT 5616  
 Db 2233 GCGGCTATTCGGGGCGCTGTGACAGATCTTCAAGGGCGCCAACTTCCGACCATCTTCT 2392  
 QY 5617 GCGCTATTCGTGATGCGGCTGTCTCCAGCGATCCGCTGGCTGGACCGCGGGCTGGCA 5676  
 Db 2293 GCGCTATTCGACGCGCGCCACCGTGACAGCAATGCTGGCGCGGAAGCTGTACCG 2452  
 QY 5677 GTTTTACGTGCTGCGCGCGGTGCGCATCTCGGTACCGCGCAATTGATCTGGCAACAG 5736  
 Db 2453 CTTTCTAGTGTCTCGCGCGGTGCGCATCTGGGCAAGGCGGACATGATGATGGAGGTGG 2512  
 QY 5737 AAGGCGCGGCTTGGCGGATGCTATTTTGTCTATCTGAGAGACACTATATGCGGGGTC 5796  
 Db 2513 AAGGCGCGGCTTATCGGACCGCATCTCGCCCATCTGGAAGAGCGGCTACTGCGGGGC 2572  
 QY 5797 TGGACAGCAGTATGTGACACACCGTATGTTTACCGCGTTGATTTTTCGCGACAGCGTGC 5856  
 Db 2573 TGGCGACAGCTGTGATCTCGCGCATCTCAACCCCTTGGCTTCCGCGACCACTGT 2632  
 QY 5857 ACGGCGATCAGGGGCTCAGCGGTTTTCGCTGCAAGCCATTTTTCGCAAGAGCGCTGTCC 5916  
 Db 2633 CCGGCGATCAGGGGCTCGGCTTCTCGGTGAAGACCGCTCTCACCGAGTCAGCGTTC 2692  
 QY 5917 GCGGCGATACCGGATGCGGATGCTATTAACCTTTATCTGTGGGGGCTGGACGATC 5976  
 Db 2693 GCGGCGATACCGGATGCTCAAGATGCTCAACCTTTATCTGTGGGGGCTGGACGATC 2752  
 QY 5977 CCGGCGGATGCTGCGGATGCTGATCGGCTCGGGAAGACGCGGCTGATGATG 6036  
 Db 2753 CCGGCGGATGCTGCGGATGCTGATGCTGCGGATGCTGCGGATGCTGCGGATGCTGCG 2812  
 QY 6037 AGGATCTGACCG-----GATGAACCAACCGCGCTGATGACGAG-----TCAC 6081  
 Db 2813 CGGATCTGCGGCTGACCGCGGATGACGCGCTGCGGCTGCGGCTGCGGATGACGAG 2872  
 QY 6082 GCAACCAATGCGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGCTC 6141  
 Db 2873 GGAAGCCATTAACAAGGCTGCAAGAGCTTCCGCGCGCGCGCTGCTGCGCGCG 2932  
 QY 6142 AACGCGCGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 6201  
 Db 2933 CATGCGGGAAGACCGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 2992  
 QY 6202 TGGTCAAGCGTGGGCGGAGCGGACGACGCGGCTGCGGATGCGAGCGCGGAT 6261  
 Db 2993 CGGCGAGGAACTGGGGGACGCGGCGGTCTCCAGCGCGCTGCGCGG-----CGAGCGCT 3049  
 QY 6262 GCGGCACTGCAATTCGAAACCGCGCGGCTTACAGCGGCTGCGGATGCAACGAC 6321  
 Db 3050 CGATGCGCTCTATGAGAGAGCCCGAACGCTTATCGCGCGCGCTTCCGCGCATCCGCG 3109  
 QY 6322 GTTTGCTGCTTTCGAGAGTGGCGCTGACGCTTCCGAGAGCTGCGGCTTTGA 6381  
 Db 3110 CTTGCGCGCTTTCGCGGAGTGGAGCGCAACGACATTTCCGAGCGCTATCCCTTCA 3169  
 QY 6382 TCAATCTGGAAGGTTTTCGATGATGCGCTGGAAGACGTTATGCTGTTTCGCGGACAC 6441  
 Db 3170 ACTGCTGGAAGGCTTTCGCGATGAGAGTGGAGGCGCGGCTACGACACCTTCGATGACAC 3229  
 QY 6442 GCTGCTTACGCTATCAGTGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6501  
 Db 3230 GCTTACCACTATGCTACACGCTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3289  
 QY 6502 CGTACGTGATGAGCGCGCTACTGATCAAGCGCTGATGTTGGGCTGCGGCTTAC 6561

Db 3290 GCGGCGGATGAGTGTGCTGAGCCGCGGACGAGACCTGAGGCTGAGCTTCCAGCTCAC 3349  
 QY 6562 CAATATGACAGGATATCTGTTAGAGACGCGGAGATGAGCTGCTGCTATCTGCGCAACAAG 6621  
 Db 3350 CAACATGCGCGGACCTGATGAGAGTGGCGCATGCGCGGATCTACCTGCGCGGCGC 3409  
 QY 6622 CTGCGTGAATGAGGCGCGGACCTGAGCGCGCGCAAGCTTGGCGATTCGCAATCGCGAGC 6681  
 Db 3410 CTTCTTTCGAGAGGCGCGGCTGCGGAGCGCGGATGCGCGCGGAGAGACCGCGCGC 3469  
 QY 6682 GCTGCGCGCGCTGCGAGCGGCTGTGTGCGGAGCGGAGCGGAGCGGATCTATGATGCGCG 6741  
 Db 3470 GGTGCGCGGCTGTGTGCGGCGGCTGCTGCGAGTGGCGGAGCGGAGCGGATCTATGACAGCGGCT 3529  
 QY 6742 CAGCGGCTGCGAGATGCGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 6801  
 Db 3530 CATCGGATGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 3589  
 QY 6802 TTACCGGGAATGCGGCTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6861  
 Db 3590 CTATCGCGCATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 3649  
 QY 6862 GCGGACCATGAAGGCGGAAAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6921  
 Db 3650 CTCACCTTCAGCGGCGGCGGAGGCGGCTTTCGCGGAGTGGCTGCGGAGGCGCTGCGC 3709  
 QY 6922 TTGCGG---CCTTGTCTATCCGAGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTG 6978  
 Db 3710 CACCGGCTTGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3769  
 QY 6979 ACAC 6982  
 Db 3770 AccC 3773

RESULT 11  
 AF218415  
 LOCUS  
 DEFINITION  
 Bradyrhizobium sp. ORS278 geranylgeranyl synthase (crtE), lycopene cyclase (crtY), phytoene desaturase (crtI), phytoene synthase (crtB), and beta-carotene ketolase (crtW) genes, complete cds.  
 AF218415  
 AF218415.1 GI:8650414  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Bradyrhizobium sp. ORS278  
 Bradyrhizobium sp. ORS278  
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.  
 REFERENCES  
 Hannibal, L., Lorquin, J., D'Orto, N.A., Garcia, N., Chaintreuil, C., Masson-Boivin, C., Dreyfus, B. and Giraud, E. Isolation and characterization of cantaxanthin biosynthesis genes from the photosynthetic bacterium Bradyrhizobium sp. strain ORS278 J. Bacteriol. 182 (13), 3850-3853 (2000)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 20309720  
 10851005  
 2 (bases 1 to 6635)  
 REFERENCES  
 Giraud, E. and Angles d'Orto, N.  
 Direct Submission  
 Submitted (02-SEP-1999) Laboratoire des Symbioses Tropicales et Méditerranéennes, TA/10J, Campus de Baillarguet, Montpellier 34398, France  
 FEATURES  
 source  
 1. 6635  
 /organism="Bradyrhizobium sp. ORS278"  
 /mol\_type="genomic DNA"  
 /strain="ORS278"  
 /db\_xref="taxon:114615"  
 /complement(273..1223)  
 /gene="crtE"  
 complement(273..1223)  
 /gene="crtE"





OY	4334	GCAGAGACGCTTTTTCCTCGCTGTTAAATCGCATGCTGTCTTCTTGCCGGAACGGCTTGAAC	43933
Dp	2386	GCCGCGCGGGCTACTATATGCTTCCGAAACGGATGCTGTCTTAAGCCCGCCAGCCCTTCGG	24455
OY	4394	AGCGCTGCGCGGTAAATGCAAGCTTTTACCGGACTGACGCGCGATGTATTAAGCCGCTTTT	44533
Dp	2446	AGCGTCAACCGAATCTGTGGCAGATTTTACGGTCTCGATCAAGCCCTGATTCGAGCGCTTCT	25055
OY	4454	AACGCGGGCGCACTGCGCCTCAGCGATTAAGCAGCGAT-----TCTGTGCGGCAAAAC	45045
Dp	2506	ACGCGCGCCGGATCAGCGCCGACAGACAAAGCTGCGCGTCTTGATGATATGTGCTGATGAAGC	25655
OY	4505	CGCCGCTCCCTCTCGGCGGAACGCGCT--GGCGCATGATGATGACTCTCCGTTACAGGG	45633
Dp	2566	CGCCGATCCCATCTCGTCCGCGCTCGCTGCTCCGAGCGACGCCCTTCAAGAACCC	26255
OY	4564	AAGAAATATATAAACGCA-----CCATATGATGTGGCGCAGGCTTCGGTGG	46095
Dp	2626	CATGAGTCAAGAGATGCAAGAACGCCAAGACAGCACTGTGATTCGGTTGCGGATTTGGTGG	26855
OY	4610	CTGCGCGCTGCGCATTCGTCTGCAAGCGGCGCGCGTGCAGCTCACTGTGAACACCG	46695
Dp	2686	GCTTTCGCTCCCATACGCTGCAATCGCGCGGATTCGCAAAACGCTCGTCGAAAGCG	27455
OY	4670	CGATTAAGCCTTGGCGGCGCGCGCTTATGTGTATCAAGATTAAGGCTTTTACTTTGATGCCGG	47295
Dp	2746	CGACAAAGTTCGCGCGCGCCGCGCTATGTCTTAAGAGAGAGACGGCTTCACTTCGACGCCGG	28055
OY	4730	TCCGACGGTGAATTAACCATCCAGCGCTATCGAGGCGCTGTTTACGCTGGCAGGCAAGCA	47895
Dp	2806	GCCCACTGTATCAACCGATCCGACTGCTGCAAGAGCTTGTTCGGCTCTCGGGGCGCA	28655
OY	4790	ACTCAGTGATATGTGCACTGTATGCGCGGTGACGCGCATTTTATCGCCTGTGCTGGAGAGA	48495
Dp	2866	GCTCGAGAACTATGTGCACTGATGCGCGGTCTCCCAATTATCAAGTGGCGCTGGAGAGA	29255
OY	4850	CGGCAAGCAGCTGAGACTGACAAACATACAGCGCGACTGGAGCAGCAGATTTGCCACTTT	49095
Dp	2926	CGGGGCCACTTTTCACTACGTCAACGACCAAGCCGAGTGAAGCGCCAGATCGCGCGTT	29855
OY	4910	TAAATCCAGGATGTGCGCGGTTACCGCGAGTTCTGGCCATTCACAGATGTGTTTGG	49695
Dp	2986	CTGTCCGCTATGTGACGCGCTATCGCGCTTCGAGTGTATAGCGAGCCCTGTGGA	30455
OY	4970	TGAGGCGCTATCTGAAACTGGGCAACCGTACTCTTTTCTGCAATTCGCGCATGCTGCGTGC	50295
Dp	3046	GGAAGGCTACGTCAAGACTGGGCAAGTGCCTGTTCCCGGATTTCCGAGACATAGTGCCTGT	31055
OY	5030	CGGCGCACAGCTGGGTGCGCTGTCAGAGCCTGGCGCAGTGTCTACAGCATGTGTGGGAATT	50895
Dp	3106	GGCGCCGCGAGCTGTGTGGGCTGTCAGAGCTATTCGACGCTCTATAGCAAGGTGTGCAATA	31655
OY	5090	TATTCATGACGATCATCTGGGCGCAGGCTTTTCCTTTACTCGTTGCTGCTGCGCGGTA	51495
Dp	3166	CGTCTCCGACAGCAATCTGCGGCAAGCCTTACAGCTTCCACTCGCTGCTGTGCGCGGCA	32255
OY	5150	TCTTTTTCGCAACGTCTTCGATCTATACCTTAATTCAGCACTGAGCGCGAATGGGCGT	52095
Dp	3226	TCCGTGGCGAAGTCTCGATCTACAGCGCTATCATGCTGCTGAGAGCGCGCTGGGAGT	32855
OY	5210	GTGCTTTCGCGCGCGCGGTACCGGTGCGCTGCT--TGATGCAATGGCGCGGCTGTTTTCG	52665
Dp	3286	CTGTGTTCCCGCGCGCGCGCACCGGCGGCTCATCAACAAAGGGGCTGTGCCAGCTGTTCAA	33455
OY	5267	CGATTTGGGCGGTGAACGTGCTGCTCAACGCGGAAGTCAAGCGCATGGAGACCGAGGGTAA	53265
Dp	3346	GGATCTCGGCGGACAGAGTGAAGCTGTGCAACCAAGCTGAGCCGGATCGAAGAGGGAACGG	34055
OY	5327	CCGATCAAGCGGTGTCCAGCTGAAGAGATGGCGCGCTTTTCCCGCGCGCGCTGTGCTC	53865
Dp	3406	GCGGGTCAAGTCCGCTGTGCTGCGGAGAGAGCGCGCTGTGCTGCGCCGACATCTGCTCGCAG	34655
OY	5387	AAATGTGACGTGTGCTACCTACGATCGCCTGTTAAGCCAGCATCTTCGCGCGCGTAA	54465

Db	3466	CAACCCGACGCTGTTTCACTTACCTTACCCGATCTCTCTGAAGACGAGCCGCTGGCGCGGCC	3525
.QY	5447	ACGCGCGGCAACGCTGAAAGCCGAAGCGGATGAGCACTCGCTGTTTGTACTTATTTTGG	5506
Db	3526	GACGCGCGCAATCTTTGATGCGCMAAGCGTTTACGATGTGCTGTTTGTGATTTTCGG	3585
QY	5507	TCTTATATATGCCCAACCCGACAGTGGCGCACCAACAGGTGTCTTTGGTCCGCGCTATCG	5566
Db	3586	CCGCGCGCGGACAGATCCGAGAGCTCAAGACCAATATCTGTCTGGCCGCGCTACCG	3645
QY	5567	TGAATGATCGATGAGATCTTCAATAGACGACCTGGCGGAAATTTCTGCGTATCT	5626
Db	3646	CGAGCTGATCAACGAGATCTTCAAGGGGCCCGCGTTGCTTGAGGATTTTCTGCTGATCT	3705
QY	5627	GCATGCGCCCTGCTCCAGACGATCCGCTGGCTGGACACCGGCGGGCTGCGGACGTTTATACGT	5686
Db	3706	GCAGCGCGCGAGCGTGAACCGATCCGCTGGCTGGCGCACAGAGCTGACGACCTATTACGT	3765
QY	5687	GCTGGCGCGGATGCCGATCTTCGATACCGCGGCAATTGACTGGCAACAGAGGCGCGCG	5746
Db	3766	GCTCTCGCGGGTGGCGCATCTTCGCGCGCGCACCATGACTGAGAGCGGTGAGGCGCGCG	3825
QY	5747	CTTGGCGGATGGCAATTTTGTCTTACTGAGAGACACATATGGCGGGTCTGCGACAGCA	5806
Db	3826	CTATGCGACCGCATCTTCGACTCTGAGAGCGCGCATTTGCGCGGGGCTGAAATGCGGA	3885
QY	5807	GTTAGTGCACACCGGTATGTTTACGCGCGTGTGATTTTCGCGACACGCTGCACGCGCATCA	5866
Db	3886	CCTGCGCACCTGCGCGCATCTTCAAGCGCGCAGAGATTTCAACACGAGCTGAATGGCATCT	3945
QY	5867	GAGCTCAGCGTTTTCGCTCGAACCATTTTGAACGCAAGCGCTGTGTTCCGGCCGCAATA	5926
Db	3946	CGGCTCCGCTTCTCGCTGAGCGGATCTTGACACAGACGCGTATTTCCGCGCCACAA	4005
QY	5927	CCGGATCCGCACTTACTACTTTTACTGTGTGGGGCTGGGACCGATCCCGGTGGCGG	5986
Db	4006	CGCGGACACAGATCAAGGGGCTGTCTGTGGGGCGCGGACCAATCCGCGGCGCGG	4065
QY	5987	TGTGCGAGGCGTGTATCGGCTCCGCGAAAGCAGCCGCGAGCTGATGTGTGAGATCTGAC	6046
Db	4066	CATTCCCGGCTGTCTGCTGGCGACAGGCGACCGCGGATCACTCTCGACGACGACGG	4125
QY	6047	CGGATGAACCAACCGCGCTGATTGACAGGTACGCGAACCATATGGCGC-----	6095
Db	4126	CGAATCGCGGAGCGCATACCTCGCGCAGCAATGACATCTGGCGCGGCTGAGCGAA	4185
QY	6096	-----AGGCTCCAAAGTTTGGCCAGCGCTAACCGGCTATTTGATCTTTCA	6142
Db	4186	GCGCTCATCCGCAACGCTTCGAAAGTTTTCGACCGCGATCAAGCTGTTCGATTTCCCGC	4245
QY	6143	ACGCGCGCATACGCTGATCTGTACGCGCTGTGTGCTCACTGTGACGATGTATGAT	6202
Db	4246	ACCGGTGCGAGCGTGCACCTGTCTTACGCGCTGTGTGCGGACATTGCGACGACGTATGAC	4305
QY	6203	GGTCAAGCGTGGGCGAAGGCGGACCGGACGCGGCTGGTGGGATGACAGGCGCGGATG	6262
Db	4306	GGACGAGATTCGGAATTCGCGCAGGCGGTGCGCGCTGGCCGCGAGATCGGGACTTTTG	4365
QY	6263	CGCCACCTGCAAAATCGAAACCCCGCGGCGCTTACAGCGGTGCCCAATGATGAACGACGG	6322
Db	4366	CAGATGTGCGCGACAGACGCGCGCAGCGCTTGAGGGGGGCGGATGCGGATCTCGGTG	4425
QY	6323	TTTGTGCTCTTTCAGAAAGTGGCGCTGACGATCAGCTTCCCGACAGCTGGCTTTTGAT	6382
Db	4426	TTTCAAGGATTTGACGCGGCTGTGACAGAGCACCGCATTTCCGACCACTACAGTGTTCAG	4485
QY	6383	CATCTGAAAGGTTTGGATGATGCGGTGAAGAACGTTATGCTGTGTTTCGGGACAGC	6442
Db	4486	CTGCTCGACGCGCTTCGCGATGGAATGTGAGGGGCGCAATATCGAGACGCTGAGCGAGCG	4545
QY	6443	CTGCGTTATCTGATCAAGTGGCGGCGTGTGGGTTAATGATGGCGCGCTGATGGGC	6502

Db	CTGAGACTACTCTATCACTGAGCGCGGTGGTCGGCGGATGATGATGGCGCATCATGGGC	4605
Oy	6503 GTAGCTGATGAGCGCGGTACTCTGATCAACCCCTGTGATTTGGGTCTGGCGTTTCAGCTTAC	6562
Db	4606 GCTCGCGGAGAGCGGACCCCTGGACCGCGCGCGCATCTCGGCATCGCGCTGCAGCTCAC	4665
Oy	6563 AATATCGCAGCGGATATGCTTGAAGGAGCGGAGAAATGGCGCGTTCATCTGCGCAAAAGC	6622
Db	4666 AACATCGCGCGTATGATGAGGATGCCAGACCGCGCCCATGATCTGCGCGAGCA	4725
Oy	6623 TGGCTGATGAGCGCGGACTGAGCGCCCGCAGCTTGGCGCAATCGCGACGCG	6682
Db	4726 TGGCTGTGCGAGCGCGCGGTGCGCGCGCGCGAGATCGCGGAAACCGCAGCATCGCGAGCG	4785
Oy	6683 CTGCGCCCCGTGGCAGCGCGTCTGTGTGCGGAGCGCGGCGTACTATCACTAGCGCGC	6742
Db	4786 GTGTTCCTGTCGTGTCGCGCGGCTGCTGATCTCGCGGAGCAGTTTACAGAGCGCAG	4845
Oy	6743 AGCGGCGTGGCAGATTTGCGCGCTCCCTTCGCGCGTGGCGATTCGCGCACCGCGCGGGCGTT	6802
Db	4846 GCAGCGCATCGCGCGCGCTGCGCGGTGTCGCGCTGGGCGGTGAGACCGCGCGGTGTC	4905
Oy	6803 TACCGGGAATTTGGCGGTAAAGTGAGCATGCGCGTGC CGCGGATGGAGATACGCGCAG	6862
Db	4906 TATGCGCAGATCGCGCGGAGAGTATGTAAGCGCGGTCCGCGCGCTGGGACCGCGCATC	4965
Oy	6863 CGCACCACTAAAGCGGAAAGCTGG	6887
Db	4966 GCCACGACAGCGCGCGCAGACTGG	4990
RESULT 12		
PMCRWTZYI		
LOCUS	6335 bp	DNA
DEFINITION	linear	BCT 15-SEP-1999
ACCESSION		
VERSION	Y15112	
KEYWORDS	Y15112.1	GI:5912291
SOURCE	Paracoccus marcusii	
ORGANISM	Paracoccus marcusii	
REFERENCE	1	
AUTHORS	Harker, M. and Hirschberg, J.	
TITLE	Carotenoid biosynthesis genes in the bacterium Paracoccus marcusii	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 6335)	
AUTHORS	Hirschberg, J.	
TITLE	Direct Submission	
JOURNAL	Submitted (12-SEP-1997) J. Hirschberg, The Hebrew Univ of Jerusalem, Dept of Genetics, Life Sciences Institute, Jerusalem 91904, ISRAEL	
FEATURES	location/Qualifiers	
source	1. 6335	
	/organism="Paracoccus marcusii"	
	/mol_type="genomic DNA"	
	/strain="MH1"	
	/db_xref="taxon:59779"	
	409..1137	
	/gene="crtW"	
	409..1137	
	/gene="crtW"	
	/function="astaxanthin biosynthesis"	
	/codon_start=1	
	/transl_table=1	
	/product="beta-carotene C-4-oxygenase (ketolase)"	
	/protein_id="CAB56059.1"	
	/db_xref="GI:5912292"	
	/db_xref="GOA:Q9RLH7"	
gene		
CDS		

gene  
CDS  
1134..1622  
/db\_xref="rEMBL:O9RLH7"  
/translation="MSAALPKALDTATSLIVSGGIIIAAMLAHVAHMLPDAANPDL  
LANAPFLTUSVGLFIANDAHGSGVPSRPAANAAGVILMTLAGSPKRCIKYK  
HNAHNRHAGTDGDDPFDHGGFVWRVARTGYTFWRGLGLPVIVTAYALLGRMMY  
VFWMLPGLILASIQLFVGTWLPKRPGHDAFPRHNRASRISDPVSLTCFHHGCVYH  
HHHHLHPVPMWRPLPSTRTKGDTA"  
/gene="crt2"  
/gene=".crt2"  
/gene="crt2"  
/function="astaxanthin biosynthesis"  
/codon\_start=1  
/transl\_table=1  
/product="carotene hydroxylase"  
/protein\_id="CAB56060.1"  
/db\_xref="GI:5912293"  
/db\_xref="GOA:O9RLH6"  
/db\_xref="rEMBL:O9RLH6"  
/translation="MTNPLIVATVLTVMELTAVSVHNMIGPLGCMGKSHHNEHDH  
ALEKNDLGLFPAVATVLTVMIMAVPLTWMILGNTVGLIYFVLDGVYHORRPP  
RIPKPKGARRLYQARHLHAVEGRDHCVSFPIYAPVDKLKDLTSGVLRNAAQOE  
RT"  
gene  
CDS  
1619..2779  
/gene="crtY"  
/gene="crtY"  
/function="astaxanthin biosynthesis"  
/codon\_start=1  
/transl\_table=1  
/product="ycopene beta-cyclase"  
/protein\_id="CAB56061.1"  
/db\_xref="GI:5912294"  
/db\_xref="rEMBL:O9RLH5"  
/translation="WTHDVLLAGALANGALIALAARAPDLRVLLDHAAGSEBGT  
WSDHSDLSPLSMIAKLPKLRANMPDOVRPPHARLATYGSLDGALADAVNRSGT  
AERDSDSIALLDAGATLSGCTRIEAGSVTDGRPHSRLLTGPHKPLGSEIETPR  
PHGVRPVIIMGITVYQRDYGITVLPKSRTRLIEDTRISDGDLDDDLAASQDY  
AAQOQWTAEGVRRERGIPLIALHDAGFMDHAEQGVGLRAGFFHPVTGYSLPYA  
AQAVAVVAGLGGPPTDALRGAIIDVYIDRARDRFLRLNRMDFRCADRRVTLLO  
RYRPHGLIRFVYAGRLSVADQLRIVTGKPIPLGYAIRCLPERPLKENA"  
2776..4281  
/gene="crtI"  
2776..4281  
/gene="crtI"  
/gene="crtI"  
/function="astaxanthin biosynthesis"  
/codon\_start=1  
/transl\_table=1  
/product="phytoene desaturase"  
/protein\_id="CAB56062.1"  
/db\_xref="GI:5912295"  
/db\_xref="GOA:O9RLH4"  
/db\_xref="rEMBL:O9RLH4"  
/translation="WNHNSPAKTAIYIGAGFGGLALAIRQSGIAITLVYARDKRG  
GAAYVHDQGVAFAGPTVITDPALKEMLALTDQMARDDVLTMPVSPFKLMPGSK  
VFDYVADNDQERQIAQENPDLEGYRFYRAEVYQGVKGTVPFLGQMLAKRA  
APALMLKAVNSYAKVATPIKDPYLRQAFYHFLVGNFSTSSIALIHALERGGY  
GVWFKGHTNOLVAGVALFERPGGQWMLAKVARIIEGARLTGVTLLADRSRLAM  
VASNDVWHTNRDLIGHTARAGOSRAKSLDRKWSMLFVLRGLREAPKIDAHHTTIE  
GRYKELVNEIIFKQPKLAEDSLYHSECTTDPMAAPGMSSTHYVLAVPHLSEAED  
WAVESPRAADTILAFLEERLIPNLRLNLTIRITVTPAFASLEIHAHGSASVSPILTL  
QSAWRRPNRDKTRINFLVGAQTHPGAGIPGVVGSAAKATVQWLSLASA"  
4278..5192  
/gene="crtB"  
4278..5192  
/gene="crtB"  
/function="astaxanthin biosynthesis"  
/codon\_start=1  
/transl\_table=1  
/product="phytoene synthase"  
/protein\_id="CAB56063.1"  
/db\_xref="GI:5912296"  
/db\_xref="GOA:O9RLH3"  
/db\_xref="rEMBL:O9RLH3"

/translation="MSDLVLTSTREITQSGSFATAKIMPRGIDDDTWMYACRHA  
DDVIDGALSGRPREAVNDPQARKLGLADTLALQCGPTTPPALRAVARRRDFQ  
AWEMDLIEGFAIDVARDYRTLDVLSVTHAGILVEMARVGVDDHPLVDACD  
GLAFOLTHIADVIDIARIGRQVGLMDIAGARVGPVSPELVYVILRLDAAEP  
YPSARVGLADIPRCAMSIIPARIVYALICLPKRGPEAYRORISTSKAKIGLLG  
IGMDVADHACRGSVGRODLTPRPHNA"  
complement (5189..6064)  
/gene="creb"  
complement (5189..6064)  
/function="astaxanthin biosynthesis"  
/gene="creb"  
/function="astaxanthin biosynthesis"  
/codon\_start=1  
/transl\_table=11  
/product="geranylgeranyl pyrophosphate synthase"  
/protein\_id="CAB56064.1"  
/db\_xref="gi:5912297"  
/db\_xref="GOA:Q9RLH2"  
/db\_xref="EMBL:Q9RLH2"  
/translation="MRDDVNPILHATLIQTRIEIAOGRGAVLOPLGMAALSGKRF  
RGMILMLAAESGVCOTIYDACAVENTHAAASLIFDLPENDAGLRPOPAITHAH  
GESRAVIGGILITEAMALLAGARAGSTVAQOLVILSRISQGLACAGDLDLHA  
KNGAVBOBODLKTGVLFIALEMIAVIKEPDABEOTQMTDFGQLRVRVPSYDLE  
VVGQOALGKDTGRDAAAPGBRRLGLAVSDIQKVSQVYASRAQLIEGVSGSKLQAE  
IEALLERVLPPYARA"

Query Match 12.2%; Score 1054.6; DB 1; Length 6335;  
Best Local Similarity 57.1%; Pred. No. 9.4e-188;  
Matches 2056; Conservative 1; Mismatches 1500; Indels 42; Gaps 6;

ORIGIN

QY 3334 CTGTACGGGGGGGGCTGATATTGTCGACGACGCGCTGTGTCAGACGACAAATCGTGGC 3393  
Db 1541 CTATGCGCCCCCGGTGCAAGCTGAAAGCAGAGACTGAAGACGTGCGGCTACTGCGGCG 1600  
QY 3394 GAAGGCCACTGATGCGCGCACTTATGATGTCATTCTGTCGTCGCGCTGCTGCTAAC 3453  
Db 1601 CGAGGCGGAGAGGCGACGTGACCCATGACGTGCTGTCGACAGGGGCGGCGCTTGCGAAC 1660  
QY 3454 GGGCTGATGGCGTGGCTTTACGCGAGCTGACGCGCGCACTTAAAGTTTCTACTGGAG 3513  
Db 1661 GGGCTGATGCGCTGCGCGCTGCGCGCGCGCGCGCGCACTGCGGCTGCTGCTGAC 1720  
QY 3514 AGTCAGGCGACGCGCGCGCGCAATCATCTGTCGTCCTTCATCGGAAAGCTCAGCGAA 3573  
Db 1721 CATGCGCGGAGACCGTCAAGAGCCATACCTGTGATGCCACGACCCGATCTGCCCG 1780  
QY 3574 GCGCAGTTTCGCTGCTGACCGCTGCTTTCGCGCGCTGCGCGCTGCTTATCAGTACGC 3633  
Db 1781 CACTGCGTGGCGCGCTGAAACCCCTGCGCGCGCAACTGCGCGACCAAGAGGTGGC 1840  
QY 3634 TTCCCACTTCGCTGCGCTGCACTGATGTGTAATTTCTGATTTGCTCGAGGATTTT 3693  
Db 1841 TTTCCCGCCATGCGCGCGCTGCGCAACCGTTTACGCGTGTGAGCGGCGCGCGCTG 1900  
QY 3694 GCCCGGCACTTACAGCAGAGTCTGCGTGGCGCGCTGACGACGCGCGCTGACGAG 3753  
Db 1901 GC---GATGCGGTGGTCCGCTCGGCGCGCGAGATCGCTGGAGACGAGCATGCGCTTG 1957  
QY 3754 GTCTCACCACCGGAGTCACTGCGAGATGCGGAGATTACAGCGCAGCGGATTT 3813  
Db 1958 CTGAGTGGCGAGGGGGCGACACTGTCGCGGCAACCGGATCGAGGCGGCTCGCTCG 2017  
QY 3814 GACGAGCGCGGCTGCGACCGCATCTGACGCTGCGCTATCAGGCAATTTGCGGT 3873  
Db 2018 GACGCGCGGCGCGCGACCGCTCCGCGCATCTGACCTTTGCTTCCAAAAATTCCTGAGT 2077  
QY 3874 CAGAGTGCAACTGCGCGCGCGCATGCGCGACGACCAATATTGATGAGCGCCAGC 3933  
Db 2078 GTCAAGATCGAGACCGACCGCCCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCA 2137  
QY 3934 GTGCATCAGACAGAGGTTATCGCTTTGTTTACACCTGCGCGCTCAGTGCAGCGCGTTTA 3993  
Db 2138 GTCAACCGAGCGGAGCGGTTATCTTATCTGCTGCGCTTCTCTCGACCGCGCATC 2197

QY 3994 CTGATTGAAGATACCACTACATCAACCATGCCAGCTGGATGCCGACAGCGCGCGCT 4053  
Db 2198 CTGATTCAGAGCAGCGCTCTATTCGATGGCGCGCATGTGACAGCAGCGCTGGCGCGC 2257  
QY 4054 CACATTCAGGATTATGCGCCACAGCGCGGCTGGAATTTGGCGCAGCTGCTGCGGAGAG 4113  
Db 2258 GCGTCCCAAGACTATGCGCCCGCGCAGCAGGTTGACCGGGGCGAGTCCGCGCGGACCG 2317  
QY 4114 CACGCTCGCTGCGCATCAGCTCAGCGGCGATATCGATGATTCTGGCAACAGCAGCAC 4173  
Db 2318 GGCATCTGCTTATTTGCTTGCGCCATGACGCGCGGCTTTCTGGCCGATCAGCGGAG 2377  
QY 4174 GGGCAACCGTGCACGCGGCTGCGCGCGCGGCTGTTTACCGCCACACCGGTTACTGCTG 4233  
Db 2378 GGGGCTGTTCCGTTGGGACCTGCGCGCGGCTTTCTTACCCCGGCTATTCGCTG 2437  
QY 4234 CCGCGCGGTTGGGCTGCGCGGAGAAATTGCGAGC---ACGCTGCCGCGCAGCTCAC 4290  
Db 2438 CCTATGCGCGCGCAGGTGGCGGACGTGTGGCGGCTGTCGCGGCGCGCGCGCGAC 2497  
QY 4291 ACGCTAGCCACTGATCGAATCCTTTGCCGTCAGACTGGCGCGAGCAGCGCTTTTC 4350  
Db 2498 GCGCTGCGCGCGCGCATCCGCGATTACGCGATGACCGGCGACGCGCGCGCTTCTG 2557  
QY 4351 CGTCTGTTAATCGCATGCTGTTCTTCCGCGACGCGCTGTAACAGCGCTGGCGCTATG 4410  
Db 2558 CGCTGCTGAACCGGATGCTGTTCCGCGCTGCGCGCGCGCGCGCTTATCCTCTCTG 2617  
QY 4411 CAGCGTTTATACCGGCTTGAACGCGGATGATTAGCGGCTTTTACGCGCGGCAACTGCGC 4470  
Db 2618 CAGCGGTTATACCGCATGCGCGCATGATGATGAAAGGTTCTATGCGCGCGCTGAGC 2677  
QY 4471 CTACGATTAAGACGCACTTCTGTGCGGCAACCGCGGTCCTCTGCGGAAAGCGCTG 4530  
Db 2678 GTGCGGATCAGCTGCCATGTCGACCGGCAAGCTCCATTCCTCTTGGCAGCGGCATA 2737  
QY 4531 CGCGATTGATGATGACTCTCGGTTACAGGGGAAAGAAATA----- 4572  
Db 2738 CGTGCCTGCCGGAACGTCCCTGCTGTAAGAAABACGATGAAACCCCATTCGCCCGCG 2797  
QY 4573 TGAACGCACTATGATGATGATGCGCAGAGCTTGGTGGCTGCGCTGCGGCAATTCGTGCG 4632  
Db 2798 CCAAAACCGCATGCTGATGCGCGGCGCAAGCTTTGGCGGCGCTGGCCATCGCGCTGC 2857  
QY 4633 AAGCGCGCGGCTGCGGCTGACGCTGCTGAAACAGCGCATTAAGCTGCGGCGCGCT 4692  
Db 2858 AGTCGCGGCGATGCGCACCACTGCTGAGGCGCCGCGACAGCCCGCGCGCGCGCT 2917  
QY 4693 ATGTGATCAAGATCAGGGTTTAACTTTGATGCGGCTCGGACGGTGAATACGATCCCA 4752  
Db 2918 ATGTCTGACAGATCAGGGGCAATGTTCTTCAAGCGGCGCGGACGATGATCAACGACCC 2977  
QY 4753 GCGCTATCAGAGCGCTGTTTACGCTGCGCAGGCAAGCACTCAGTATTAATGTCAGCTGA 4812  
Db 2978 ATGCGCTCAAGAGCTGTGGCGCTGACCGGGCAGGACATGGCGCGGAGGTGAACGCTGA 3037  
QY 4813 TGCCGTTAGCGCAATTTATGCTGCTGTGGGAAAGCGGAGGCACTGATCAGACA 4872  
Db 3038 TGCCGTTGTGCGCTTCTATCGCTGATGTGGCGCGGCGGAAAGTTCGATTAACGTGA 3097  
QY 4873 ACAATCAGGGCGCGCTGAGAGAGAGATGCACTTTTATCCCAAGATGTGCGCGGTT 4932  
Db 3098 ACGAGGCGGATCAGCTGGAAGCCAAATCGCCAGTTCAACCGGAGACTGGAAGGAT 3157  
QY 4933 ACCGCAAGTTTCTGCGCTATTTACAGAGATGTGTTTCGTGAGGCTATCTGAAACTGCGCA 4992  
Db 3158 ACGTGCCTTCTCGATTTAAGCTGAGGAGGTATCAGAGGGGCTAAGTCAAGTGGGCA 3217  
QY 4993 CCGTACTTTTCTGCAATTTCCGGAATGCTGCGTGGCGCGGCGCACAGCTGGGTGCGTGC 5052  
Db 3218 CCGTGCCTTCTCTCAAGCTGAGCCAAATGCTCAAGAGCGCGCGCGCGCTGATGAAGCTTG 3277

QY	5053	AGGCTGCGCAGTGTCTCAACGAGTGGGGAAATTATTCAGACATATCTGGCC	5112
Db	3278	AGGCTTACAAATTCGTCCATCCAAAGTCGACCTTCATAAAGAACCTATCTGGCC	3337
QY	5113	AGGCTTTTCCCTTTCACCTGTGCTGGTCGGCGGTAATCCTTTTGCAACGTCCTCGATCT	5172
Db	3338	AGGCTTTTTCGTATCAACGCTGCTGGTGGCGGGAAATCCCTTCTCGACAGCTCGATCT	3397
QY	5173	ATACCTTAATTCACGCACTGGAAGCGGCAATGGGCGTGTGTTTTCCGCGCGGCGTACCG	5232
Db	3398	ATGCGCTGATCCACGCGCTGGAGCGCGCGCGGGGTCTGGTTGCCAAAGGCGGACCA	3457
QY	5233	GTGGCTGTGTGATGCGCATGCGCGGCTGTTTTCGCAATTTGGGGGTGTAATCTGCTGCTCA	5292
Db	3458	ACCAGCTGTGTGCGGGCAATGGTCCGCTGTTTCGAACCGGCTGGGGCCAGATGATGCTGA	3517
QY	5293	ACGCCGAAGTCAGCCAGCTGGAGACCGAGGGTTAACCGCATAGCGGTGTCAGCTGAAG	5352
Db	3518	ACGCCAAGTGTGCCCGGATCGAGACCAAGGCGCGGACCAACGGGCTCACCTTGGCG	3577
QY	5353	ATGGGCGCCGTTTTGCGCGCGCGCGCGCTTGGCTAAATGCTAGCTGTGTGATACCTAG	5412
Db	3578	ACGGGCGGTCTTTTGAGGGCGACATAGTCCGCAAGACGGGAGTCAATGCAACAATATC	3637
QY	5413	ATCCGCTGTAAAGCCAGATCTCTCGCGCGCTAAACGGCGGGCAACGCTGAAGGGCAAC	5472
Db	3638	GCGACCTGTGTGGGCAACAGCGCGCGGGCGAGACCGCGCCAAATCGCTGACCGCAAGC	3697
QY	5473	GGATGAGCAATCGCTGTTTTGTACTCTATTTTGGCTTAATCATGCCACCGGAGCTG	5532
Db	3698	GCTGTGCTATGCTGTTGTCGTGCTGCATTTCCGATCTGCGGAGGCGCCCAAGACATCG	3757
QY	5533	CGCACCAACAGGTGTGCTTTGTCGCGCTATCTGTGAATTGATGATGATCTTTCATA	5592
Db	3758	CGCATCAACATCTGTTGCGGCCCGCGCTACAAAGAGCTGTCAAGATATTCAAG	3817
QY	5593	GCAACCACTGTGGCGAAGATTTCTGCTGTATCTGACATGCGCCCTGCTCCAGCATCGT	5652
Db	3818	GCCGAAGCTGGCGAGATTTCTGCTGTATCTGCAATTCGCTCCGCAACCCATCCGG	3877
QY	5653	CGCTGGCACCGGGGGGCTGGGCGAGTTTACGTGCTGGCGCGCGTCCGATCTCGTA	5712
Db	3878	ACATGAGGCTCTCCGGGCAATGCCAATACGATACGTCGCGCCCGGTGCGGATCTGAGCC	3937
QY	5713	CCGCCGAATTTGACTGTGCAACAGGAAGGGCCGCGCTTGGCGGATCGATCTTTTCTATTC	5772
Db	3938	GCGCGAGATGATTTGGGCGGTCGAGGGGGCGCGCTATGCCGACCGATCTGGCATTTCC	3997
QY	5773	TGAGGAGCACTATATGCGGGGTCTGCGACAGCACTTATGTAACAACCGTATGTTTACGC	5832
Db	3998	TGGAGGAGCGGCTGATCCCGAACTTGGCGCGCACTGACCCAGACCGGCAATTTTCAGC	4057
QY	5833	CGTTGATTTTTCGGGACACGCTGACACGGCATCAAGGCTCAGCGTTTTCGTTCGAACCA	5892
Db	4058	CCGCGAATTTTCGCAGTGAATGAACGCCCATCAACGGAAGGCTTTTTCGTTCAGCGCA	4117
QY	5893	TTTTGACGAAAGCGCTGTTTCCGGCGCAATAACCGGATGCGCAATTTACTAACCTTT	5952
Db	4118	TCTCGAGCAATCCGCGTGTGTTCCGGCGCGCAACCGGACAAACAGATCCGCAACTTCT	4177
QY	5953	ATCTGTGTGGGGCTGCGACCGCATCCGGGTGCGGTGTGCGAGGCGTATCGGCTCCGCGA	6012
Db	4178	ATCTGTGTGGGGCGGGCAACCATCCGGGCGGGGATTTCCGGGGGTGTGGGGCTCGGCA	4237
QY	6013	AAGGGAACCGCCACTGTATGTTGGAAGATCTGACCCGATGAACCAACCGCGCTGATTTGA	6072
Db	4238	AGGCAACGGCCCAAGTATGCTGTCCACCTGG---GCAAGGCAATGACGATTTGGTCT	4294
QY	6073	GCAAGTACGCAAAACCAATGAGCGAGGGCTCAAAAGTTTCCGCAAGCTACCCGGCTATT	6132
Db	4295	GACCTCGACCGAAGCGATCAACCCAAAGGTGCGAAAGCTTTGTCCACGGCGGCAAGCTGAT	4354
QY	6133	TGATCTTTCACGCGCGCGAGTACGCTGATGCTGTACGCTGATGCTGATCACTGTGAAGA	6192

[illegible]



AUTHORS Misawa, N., Satomi, Y., Kondo, K., Yokoyama, A., Kajiwara, S., Satto, T., Ohtani, T. and Miki, W.  
TITLE Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway  
JOURNAL J. Bacteriol. 177 (22), 6575-6584 (1995)  
MEDLINE 96062243  
PUBMED 7592436  
3 (bases 1 to 5373)  
REFERENCE Misawa, N.  
TITLE Direct Submission  
Submitted (23-JUN-1995) Northhiko Misawa, Marine Biotechnology Institute; 3-75-1, Heita, Kamaishi-shi, Iwate 026-0001, Japan  
(E-mail:northhiko.misawa@bio.jp, Tel:81-193-26-6581, Fax:81-193-26-6584)  
FEATURES  
source Location/Qualifiers  
1..5373  
/organism="Paracoccus sp. MBIC1143"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:81397"  
/note="Agrobacterium aurantiacum is distributed from Marine Biotechnology Institute as Paracoccus sp. MBIC1143. Its 16S rRNA sequence is shown in accession AB008114.-cloning vector: pAK32"  
139..867  
/gene="crtW"  
139..867  
/gene="crtW"  
/citation=[2]  
/codon\_start=1  
/transl\_table=1  
/product="beta-Carotene ketolase"  
/protein\_id="BAA09591.1"  
/db\_xref="GI:1136631"  
/translation="MSAALPDKADLTATSLVSGIIAAMLAHVAHMFDAADHPI LAIANTGLTWLSVGLFIADAHGSGVPRPRANAMQVLMVLAGFSWRMTYK MAHHRAGTDDDDPDHGGFVRWYAFIGYFEMRGGLLPVITVYALILGPRWY VFWPLPSILASIQLFVGTWLPHPGHDAPRPHNARSSISIPVSLNCFHGGYH HEHMLHPVPMWRLPSTRTKQDTA"  
864..1352  
/gene="crtZ"  
864..1352  
/gene="crtZ"  
/citation=[2]  
/codon\_start=1  
/transl\_table=1  
/product="beta-Carotene hydroxylase"  
/protein\_id="BAA09592.1"  
/db\_xref="GI:1136632"  
/translation="MTNPLIVATVLMELTAYSVHRMIMHGPLGGMKSHHEBDH ALEKNDLYGLVFAVATVLTFTVGMWMAVLMWIALGMVTYGLIFVLHDGLVHQRMR RYIPRKGYARRLYQAHRLHNAHVEGRDHCVSFGFIYAPVDKLDKMSGLVRAEAG RT"  
1349..2509  
/gene="crtY"  
1349..2509  
/gene="crtY"  
/citation=[2]  
/codon\_start=1  
/transl\_table=1  
/product="Lycopene cyclase"  
/protein\_id="BAA09593.1"  
/db\_xref="GI:1136633"  
/translation="WTHDVLVAGLNGLIATLALRARPDLRLVLDHAGSDGRT WSCHPDLSPLMLRKLPLRANPDDVPRPHARLRTGSLDAALADAVVRSG AEIRMDSIDALDAQATLSCGTRILEAGAVDGRGAQPSRLTVGPKFVELETRR PHGVPRLMDATVTOODGYFIYLLPFSPTRIILETRYSDDGLDDDLAAASHYR AROOGTGAEBVRERGLIPILAHDAAGFADHAGAPVVLRAGFPHPLVGYLPPA AOVADVAGSGPCTGALRGALRDVALIDRARRDPLRLINRMIFRCGADPRRTLLQ RFTMPHGLIRFPAAGRLSVADQLRITVTKRPITPLGTAIRCLPRPLIKENA"  
2506..4011  
/gene="crtI"  
2506..4011  
/gene="crtI"  
CDS

ORIGIN  
Query Match 12.1%; Score 1053.6; DB 1; Length 5373;  
Best Local Similarity 57.7%; Pred. No. 1,5e-187;  
Matches 2079; Conservative 0; Mismatches 1469; Indels 54; Gaps 9;  
gene  
CDS  
4008..4913  
/gene="crtB"  
4008..4913  
/gene="crtB"  
/db\_xref="GI:1136634"  
/product="Phytoene desaturase"  
/protein\_id="BAA09594.1"  
/translation="NNASHPAAKTAIVIGAGFGGLALAIRLQSGAGIATTLVBARDKRG GRAYVNDQDGLFADGPTVITDPPDALKEMLALGDMARDVTLMPSPEYRLMPPGSK VPVYVNEADOLEROIACQFNPDDEGRYRPRDVAEEYVGGYVLGTVPFLKLGOMLKA APALMLKAYKSHAKVATFIKDPYLRQAFSHYITLVGSGNPSTSYALNHLERGG GWAFAGAGTQVLAVGVALFERDGGQMLNAKVARDTGPRTGTVTLADRAHTLMD VASNGVMNRYDLGHTARQSGRASKLAKRMSLFVHLFGLRABPDVAHTTLF GPRYKLVNIEIFGPIIADPFLSYLHSPCTTPBEMAPGMSHTYVLAAPVHLORADID MAYEGRYVADRILASLEERLIPVLRANLFTTRFTSPDSELANHSGASVSEPIIT QSAWPRPHNRDKTIRNFYLVGACTHGGAIIPGVYSKAKTAQVMSDLASA"  
3334 CTGAGCGCGCGCGCTGATTTGTGCGAGCGGCGCTGTGCACGACGAATCTGCTGCG 3393  
1271 CTATGCGCCCCCGGTGCAAGCTGAGCGAGCACTGMAAGATTCGGGGCTGTCGGGCG 1330  
3394 GGAAGCCACCTGATGCGCGGACCTTATGATGTCATTCTGGTCGTGCGGCTGCGCTAAC 3453  
1331 CGAGCGCGAGAGCGCGCATGACCATGACGTGCTGCGCAGGCGCGGCTTGGCAAC 1390  
3454 GGGCTGATTCGGCTGCGCTTTACCGCGACGCTGCGACCCGCACTTAAAGTTTGTACTAGAG 3513  
1391 GGGCTGATTCGCTGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGGAC 1450  
3514 AGTCAGGCGGAGCGGCGCGGCAATCATACCTGCTGTTCCATCGCGAAGAGCTCAGCGAA 3573  
1451 CATGCGCGAGAGCGCTGACAGCGGCACTGATCTCTGCCACAGACCCCGACCTGTGCGCG 1510  
3574 GCGCAGTTTGTGCTGCGCTGAGCGCGCTGCTTGGCGCGCGCTGCGCGGTTATCAGGTACG 3633  
1511 GACTGCGTGGCGGCGCTGAAGCCCTGCGCGCGCGCAACTGAGCGCCGACGAGAGTGGCG 1570  
3634 TTGCCCCACCTGCGCTGCGCGAGCTGATGATGATTTGCTTCANTTGCTCGGAGGANTTT 3693  
1571 TTTCGCCCGCATGCGCGCGCGCTGGCCACCGGTTACGGGTGCTGAGCGGGGCGCGCTG 1630  
3694 GCCCGGCACTTACAGCAGGTGCTCGGTCGCGCGCGCTAGCAGCCGAGCGCGGTACGGAG 3753  
1631 GC--GGATCGGTTGTCGCGTGGGGCGCGGAATCGCTGGGACAGCGACATCGCCCTG 1687  
3754 GTCTCACCCACCGGCGGTCAAGCTGCGCGATGCGCGGATGTTTACAGCGGACGCGGTGATT 3813  
1688 CTGAGTCGAGAGGGGCGAGCGCTGCTCGCGGACCCGAGATCAGAGGCGGCGCGGTCTTG 1747  
3814 GACGAGCGCGGCGCTGCGAGCGCGACACCGCATCTGAGTCCGCTATCAGGATTTGTCGGT 3873  
1748 GACGGGCGGGCGCGCGCGCTGCGCGCATCTGACCGTGGGTTTCCAGAAATTCGTGGGT 1807

QY 3874 CAGAGTGGCAATGTGCGCGCGGCATGCGCTGCAGACGCAATATGTATGACGCGACG 3933  
DB 1808 GTGGAATCGAACCGACCGCCCCCAAGGTGGTCCCGCCGATGATCAATGACGCGACCC 1867  
QY 3934 GTGCATCAGCAGAGGGTTATCGCTTGTATTAACCTCGCGCTCAGTGCACACCGCTTGA 3993  
DB 1868 GTACCCAGAGAGACGGGTACCGCTTCACTATCTGTGTGCTCTCTCTCCGACGGCGATC 1927  
QY 3994 CTGATTGAAGATACCACTACATCAACCATATGCCATCGCTGATGCGCAAGCGCGCGCT 4053  
DB 1928 CTGATGAGAGACAGCGCTATTCGATGCGGCGATGTGACGACGCGCTGCGCGCG 1987  
QY 4054 CACATTACGATATATGCCCAACGCGCGGTGGAATTTGCGCACGCTGCTCGCGAGAG 4113  
DB 1988 GCGTCCCACTATATGCGCGCGACAGGCGTGAACCGGGCCGAGGTCGCGCGC---GAA 2044  
QY 4114 CACGCGTCTGCGCATCAAGCTCAGCGCGCATATCGATCAGTTCTGG---CAACAGCAG 4170  
DB 2045 CCGGCACTCTTCCCATTCGCGCTGCGCCATGATGCGCGGGCTTCTGGCGCATCAGCG 2104  
QY 4171 CACGCGCAACCTGACAGCGGCGTGGCGCGGACCTGTTTACGCGCAACCGGTTACTCG 4230  
DB 2105 GCGGGGCTGTTCCCGTGGGACTGCGCGCGGGTCTTTCATCGGTCACCGGCTATTG 2164  
QY 4231 CTGCGCGCGCGGTGGCGCTGCGCGGAGAAATTTGCG---CAGCGTCCCGCGACGCT 4287  
DB 2165 CTGCGCTATGCGCAAGGTGGCGGAGGTGTGGCGGCTGTCTCGGGCGCGCGCGCAC 2224  
QY 4288 CACAGCTGAGCACTGATCGAATCCTTTCGCGCTGAGCATGCGCGCGACGCGCTT 4347  
DB 2225 GAGCGCTGCGCGCGCATTCGCGATTAACGATCGACCGGGCGCGCGCGCATCGCTT 2284  
QY 4348 TTCGCTGTAAATCGCATGCTGTTCTTTCGCGAGCGCTGAAACGCGTGGCGGTA 4407  
DB 2285 CTGCGCTTGTAAACGCGATGCTGTTCCGCGCTGCGCGCGCGCATATACCTG 2344  
QY 4408 ATGCAAGCTTTTATCGCGCTTGAACGCGGATTTATGACCGCTTTTACCGCGCAATG 4467  
DB 2345 CTGCAAGGCTTTCACCGCATGCGCATGACATGATCGAACGCTTCTATGCGCGCGCTG 2404  
QY 4468 CGGCTCAGCATTAAGACGCAATTTGTGGGGCAACCGCGCGCTCGCTCGGGAAGCG 4527  
DB 2405 AGGCTGCGGATCAGCTGCGCATGTAACGCGCAAGCTCCCATCTCTTGGACGCGC 2464  
QY 4528 CTGCGCGCATTGAATGATGACCTCTCCGTTTACAGGAGAAATTA----- 4572  
DB 2465 ATCGCGCTGCGCGCAAGCTCCGCTGAGAGAAACGATGAAACGCGCATTCGCGC 2524  
QY 4573 ---TGAAACGCACTATGATGATGGCGAGGCTTCGCTGCGCTGCGCGATTCGTC 4629  
DB 2525 CGGCGCAAGACCGCGCATGATGCGCGCAAGCTTTCGCGCGCTGCGCATTCGCGC 2584  
QY 4630 TGAAGCGCGCGCGGTGCGCGTCAAGCTGTAACGCGCGCATTAACCTGCGCGCGCG 4689  
DB 2585 TGAAGTCCGCGGATGCAACACCTGTGTCAAGCGCGCGCAACCGCGCGCGCGCG 2644  
QY 4690 CTTATGTATCAGATCAGAGGTTTATCTTGTATGCGCGTTCGCGCGGATTAACGATC 4749  
DB 2645 CTTATGTCTGGACGATCAGAGGCGATCTCTTCAACGCGCGCGCGCGCATCAACGAC 2704  
QY 4750 CCAAGCTATCAGAGCGCTGTTTACGCTGCGAGGACCACTCAATGATTAATGTGAC 4809  
DB 2705 CCGATGCGCTGAAGAGCTGTGGCGCTGACCGGGACGACATGACGCGGACGTCGCG 2764  
QY 4810 TGAATGCGGATGAGCGCATTTTATGCGCTGATGCGGAGAGCGAGGCGACGTCGAC 4869  
DB 2765 TGAATGCGGATGAGCGCATTTTATGCGCTGATGCGGAGAGCGAGGCGACGTCGAC 2824  
QY 4870 ACAACATCAGGCGCAGCTGAGCAGCATGTCACATTTTAAATCCCAAGATTCGCGC 4929  
DB 2825 TGAACGAGCGCGATCAGCTGGAACGCAATCGCGCATTCACACCGGACGACCTGAGG 2884

QY 4930 GTTACCGGCACTTCTGCGCTATTCACAGATATGTTTCTGTAGGGCTATCTGAACCTGG 4989  
DB 2885 GATACCGCGCTTCTCGTATTAAGCTGAGAGGTGTACAGAGGGCTACGTAAGCTGG 2944  
QY 4990 GCAACGTAACCTTCTGCAATTTCCGACATGCTGCGTCCGCGGACAGCTGGTGGC 5049  
DB 2945 GCAACGTCCTTCTCAAGCTGGCGAGATGCTCAAGCGCGCGCGCTGATGAAGT 3004  
QY 5050 TGAAGCTTGGCGCAATGCTACAGATGTCGCGCAATTTATTCATGACATCATGTC 5109  
DB 3005 TGAAGCTTGAAGATGCTACAGATGTCGCGCAATTTATTCATGACATCATGTC 3064  
QY 5110 GCGAGCTTTCCTTCACTGCTGCTGCGCGGATTCCTTTTGAACGCTTTCGA 5169  
DB 3065 GCGAGCTTTCCTTCACTGCTGCTGCGCGGATTCCTTTTGAACGCTTTCGA 3124  
QY 5170 TCTATACCTTAAATTCACGCACTGAGCGCGAATGAGGCGTGTTCGCGCGCGTGA 5229  
DB 3125 TCTATGCGCTGAACGCGCTGAGCGCGCGCGGCTGTGCTTCGCGCGCGGCA 3184  
QY 5230 CCGGTGCGCTGCTGATGCGATGCGCGCGCTGTTTTCGATTTGGCGGTAACGTCGTC 5289  
DB 3185 CCAACCGCTGCTGCTGCGCGCGCTGCTGCTGCGCGCTGCGCGCGCGCGCGCG 3244  
QY 5290 TCAACGCGAAGTCAAGCTGAGACCGAGGCTAACCGCATCAGCGGTGTCAGCTGA 5349  
DB 3245 TAAACGCGAAGTCAAGCTGAGACCGAGGCTAACCGCATCAGCGGTGTCAGCTGA 5399  
QY 5350 AGGATGCGCGCTGTTTTCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5409  
DB 3305 CCGAGCGCGCGCTTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3364  
QY 5410 AGCATGCTGTTTAAACGCACTGTCGCGCGCTTAAACGCGCGCGCGCGCGCGCGCG 5469  
DB 3365 ATGCGCATGCTGCG 3424  
QY 5470 AGCGATGAGCACTGCGCTGTTTGAATCTATTTGCTGCTATCATGCGCGCGCGCG 5529  
DB 3425 AGCGATGAGCACTGCGCTGTTTGAATCTATTTGCTGCTATCATGCGCGCGCGCG 3484  
QY 5530 TGGCGACCAACGAGTGTGCTTGTGCTGCGCGCTATGCTGATGATGATGATCTTCA 5589  
DB 3485 TGGCGATCAACCACTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3544  
QY 5590 ATAGCAACGAGCTGCGGAGATTTCTGCTGATGTGATGCGCGCTGCTGCGCGATC 5649  
DB 3545 AGGCGCGGAGCTGCGGAGATTTCTGCTGATGTGATGCGCGCTGCTGCGCGATC 3604  
QY 5650 GCTGCTGCGACCG 5709  
DB 3605 CCGAGATGCGCGCTGCGGCGATGTCACGCGATTAATGCTGCGCGCGCGCGCGCG 3664  
QY 5710 GTACCGCGCAATTTGCTGCGCAACAGAGAGCGCGCTTGCAGATGCGATTTTCTT 5769  
DB 3665 GCGCGCGGAGATTTGATTTGGCGCGTGAAGGCGCGCGCTATGCGCGCGATCTGCG 3724  
QY 5770 ATCTGAGAGACCTATATGCGCGGCTGCGACAGAGATTAAGTGAACACCGTATGTTA 5829  
DB 3725 GCTTGAAGAGCGGCTGATTCGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 3784  
QY 5830 CCGCGCTTGAATTTTTCGCGACCGCTGACCGCGATGAGGCTCAAGGCTTTTTCGCTG 5889  
DB 3785 CCGCGCTTGAATTTTTCGCGACCGCTGACCGCGATGAGGCTTTTTCGCTGAGC 3844  
QY 5890 CCAATTTTGAACGCAAGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5949  
DB 3845 CCAATTTTGAACGCAAGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3904  
QY 5950 TTTATCTGATGAGGCTGCGACGATCCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 6009  
DB 3905 TTTATCTGATGAGGCTGCGACGATCCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 3964  
QY 6010 CGAAGCGACCG 6069



complement (3368. .4516)  
/gene="city"  
/codon\_start=1  
/transl\_table=1  
/product="lycopenene cyclase"  
/protein\_id="AAC44851.1"  
/db\_xref="GI:1842245"  
/translation="MSHDLILAGAGLSGALLALAVDRRPARILVMDARSGPSDQHT  
WSCHTQSLSPENWLAISPIRGEWTDQEVAPPDHSRLRTTYGSIIEAGALIETDQAP  
LWMNTHVALSDIDTGATLIDGSRILEAAVCIDRAGVETPLTVGQKFGVIEIADGV  
GGERPMIMATVPOWDQYRFLYLPESFRILIEETRYSGDGLDDGALQASIDVAA  
RGGWQEMRRERRGTLIPALAHADIGFPRDHQAQVAPVGLAGLEPHVTGSLPYAAQ  
VADALAAARULTTASARRAVRGALIDRRARDPFLRLNMLERGCEPPDRRLTLORFYR  
LQPIELERYAGRLTLADRLRITVGRPPIPLSQAVRCLPERPLQERA"  
complement (4513. .5022)  
/gene="crtZ"  
complement (4513. .5022)  
/gene="crtZ"  
/gene="crtZ"  
/codon\_start=1  
/transl\_table=1  
/product="b-carotene hydroxylase"  
/protein\_id="AAC44852.1"  
/db\_xref="GI:1842246"  
/translation="MSTAAAILTVLITTAANEELTVSYHRVIMHIGPLGCMGHNKSHDE  
DHIDALEKNDLGVIFAVISIVLFAIGMGSDLAWMLAVGTGGLIYFLHDELVHGVL  
RMPFVYVPRGTLRRVYQAHNRHFAHGRENCVSGFTWAPSDVSLKAEILKRSGLALK  
DREGADRNT"

## ORIGIN

Query Match	11.9%;	Score 1032.2;	DB 1;	Length 5188;
Best Local Similarity	57.1%;	Pred. No. 1.6e-183;		
Matches 2031;	Conservative	0;	Mismatches 1493;	Indels 33;
				Gaps 7;

OY	3419	ATTGATGTCTATCTGGTCCGTGCTGCGGCCCTGGCTTAAACGGGCTGATATTCGGCTGTCTTAAAGCC	3478
Db	4509	ATGATCTGCTGATTCGGGCGCGGGGGGCTGTCCGGTCCGTATATCCGCTTGGCCGTTTCCG	4455
OY	3479	AGCTGACACCCGCACTTAAAGTTTGTCTACTGAGATGCATGACCGGCGAGCCGGCCGGCAATC	3538
Db	4449	ACCCGAGACCCGGATGCGCGGCATCGTGAATGCTGACGCGCGGTCCGGCCCTTGGACCAAGC	4399
OY	3539	ATACCTGGTCTGTTCCATCGCGAAGACGTCAAGCAAGCCGATTTTCGTGCTGACCGGC	3598
Db	4389	ACACTGGTCTCGCCACAGACAACGATCTTTTGGCCGGAATGCGTGGCGGGCCGTGTGCCCA	4333
OY	3589	TGCTTTCCGGCGCGCTGCGCCCGGTTATCAAGTACGCTTCCCCACCTCGGTGGCCAGCTGG	3655
Db	4329	TTTCGTGCGGGGAAATGAGACGGAATCAGAGAGTCCGCTTTCCGACCATTTGCGCCCGCTGA	4270
OY	3659	ATGGTGAAATATTGCTCGATTCGCTCGGAGAGATTTCGCCCGGACCTTAAGACAGATGCTCG	3718
Db	4289	CGACAGGCTTAAGCTCGATTCGAGCGGGGGCGGCTGATGGGTGCTGAG-----G	4219
OY	3719	GTGCCGCGCTTACGCAACCCGACGCGCGGTCAAGGAGTGTCAACCCACGGGGGTCAAGCTGG	3778
Db	4218	GTGTGTGATCTCGGTGGAATATACGATATGTCCGACCGCTGAGACGATATCCGGCGGAGCGCTGA	4155
OY	3779	CGGATGCGGGGATTTTACAGGCGCAGGCGGTGATTTGACGAGACGCGGGCTGCAGCCGACAC	3838
Db	4158	CGGACCGGCTCCGGGATTCGAGGCTGCTCGTGAATGACGCGCCGTGTGTCCGTCGAGAACCC	4099
OY	3839	CGCATCTGCGACCTGGCTTATCAGAGGATTTTCGTCCAGAGTGGGAACTGGCCGGGCGCG	3898
Db	4098	CGCACTGACCGGTGGGTTTCCAGAAATTCGTGGGGGTGAGATCGAGACCGACGCGCCCCC	4033
OY	3899	ATGGCTGTGACAGGCAATATTGATGAGCGCAGCGCTGATCAGCAGGAGGTTATCGCT	3955
Db	4038	ATGGGTGTGAAGGCCCGCATGATCATGAGAGCGGACCGTTCCGAGATGGAAGGGGTATCCGCT	3979
OY	3959	TTGTTTACACCTGTGCGCTCAAGTCCAGCCGTTTACTGATTTGAGATGACCATCAATCA	4018
Db	3978	TCATCTATCTGTGCTTTCAGTCCCAACCGGATCTTATCGAGATATCGGCTTACAGC	3919

QY	4019	TCATATCCAGCCTTGGATATGCGCAACAGCGCGCGCCCTCAATTACGGATTATAGCCACAGC	4078
Db	3918	ACGGCGGGGATCTGTGAGATGAGCGCGCTGGCGCAGGCGCTGCTGGACTATATGCGCCAGCG	3859
QY	4079	GGCGCTGGAAATTTTCGCGCAGCTGTCTGGCGAGAGACAGCGCTCGCTCCGATCCACGCTCA	4138
Db	3858	GGGCGCTGGATCCGGGCGAGAGATATGCGCGCGGAAAGGGGCAATCTGCCATCGCGCTGGCGCC	3799
QY	4139	GGCGCGATATCGATCAGTTCTGGCAACAGCAGCAGCGGCAACCGTGCAGCGGCGTGC	4198
Db	3798	ATGACGCCCATATGCTTCTTGGCGCGACACAGCGCGCAGGGGGCGGTGTCGGATTGGCTGGGGGG	3739
QY	4199	CCGCACTGTTTACAGCCCAACACCGGTTATCTGCTGCTCCCGCGCGGTAGCGCTGGCGGAGA	4258
Db	3738	CAGGCGTGTTCACACCGCGTCAACCGGATATTCGCTGCCCTATGCGCGCACAGGTCCGGATG	3679
QY	4259	AGATTGCGACGAGCTGCGCGCGACGCGTCAACAGCTGACGACATCGATCGAATCTTTTG	4318
Db	3678	CCATTCG--CGGGCGCGACCTGACACACCGCGCTCGCCCGTGCCTGG--TGCGCGGCTGGG	3622
QY	4319	CCCGTACGACATGCGCGGAGCAGCGCTTTTCCGTCGTTAAATCGCATGCTGTTCTTG	4378
Db	3621	CCATCGATTCGGCGGGATATCGGACCGGCTTCTCGCGCGCTGGAACCGGATGCTGTTCCGG	3562
QY	4379	CCGGAACGGCTGAACAGCGCTGCGCGCTAATGCAAGCGTTTATCCGCGCTGAACCGCGAT	4438
Db	3561	GCTCCCGCGCGACCGTGTGCTATGCTGCTGTCAGCGGTTCTACCGGCTTGCAGACGCG	3502
QY	4439	TGATTACCGCTTTTAAGCGCGGGAACGCGCCCTCAGCGAATMAAGCAGCATTCGTGGG	4498
Db	3501	TGATTCGAGCGCTTTATATGCGGGGGCGCTGCACATTGCGCAGCGGCTTGCATCGTCAACG	3442
QY	4499	GCAAAACGCGCGGTCCTCTCGGCGAAGCGGCTGCGGCAATTGATGATGACTCTTCGCTTAC	4558
Db	3441	GAGCCCGCGCCATTCGCGCTGTGCGAGCGCGTGGCTGCTGCCGAAGCGCCCTGCTGCG	3382
QY	4559	CAGGGAAGAAATTAATGAAGCAGCACTATGTGATTGGCGCAGGCTTCGTTGGCTGCGCT	4618
Db	3381	--AGGAGAGAGCATGATTTCCGCGCATTCGTCATTCGGCGCAGGTTTCGGCGGGCTTGGCGCT	3325
QY	4619	GGCGATTCGTCGTGGAACCGGCGCGGCGTGGCGGTACGCTGTGGAACAAGCGCGATTAGCC	4678
Db	3324	TGCCATCCGCGCTGGAATTCGGCGCGGCAATCGGACCAACATGTCAGAGCGCGCGCAAGCC	3265
QY	4679	TGGCGGGGCGCCCTATGTGATTCAGAGTCAAGAGGTTTACCTTTATGCGCGTCCGACGCT	4738
Db	3264	CGGCGCGCGCGCTATATGTCTGGAACATACAGGCGCAACGTCCTTCGATACAGGCCCTGACGCT	3205
QY	4739	GATTACCGATCCACGCGCTATTCGAGGCGGCTGTTTACGCTGGCAGGCAAGCAACTCAGTGA	4798
Db	3204	CGTACCGGACCCCGACAGCCTCGAGAGAGCTGTGGGCCCTCAGCGGCGCAACGATGAGAGG	3145
QY	4799	TTATGTGACTGATGCCGCTGACGCCATTTTATCGCTGTGCTGGGAAGAAGCGCAGCA	4858
Db	3144	TGAGGTGACGCTGTGCGCGGTCTCGCCCTTCTACCGGCTACATGCGCGGACGCGCCAG	3085
QY	4859	GCTGGACGACGAAACAATCAGAGCGGCGGCTGGAGACAGATGTCACATTTTATCCCA	4918
Db	3084	CTTGAAATATGCTGAACACACGACGAGCTGTGATCCGCAAGGTCGCTCTTCAATCCGCG	3025
QY	4919	GGATGTGCGCGGTATCCGCGAGTTTCTGAGCTTATTCACAGGATGTGTTTGTGAGGGCTA	4978
Db	3024	CGATGTGATATGCTATATGCGCGCTTCCAGATATACCGGAGAGGTCTATATGCGAGGGGTA	2965
QY	4979	TCGAAAACCTGGGCAACGTAACCTTTTCTGCAATTCGCGCAATGTCTGTCGCGGCGACA	5038
Db	2964	TCGTAACTGGGGACACAGCCCTTCTGAAGCTGGGCGCAATGTCTGAACCGCGCGCGCG	2905
QY	5039	GCTGGGTTCGCTGAGGCGCTGGCGCAATGTCTACAGATGTGTGGCGAAATTTTATTCATGA	5098
Db	2904	GCTATATGCGCTTGAGGCAATACCGCTCGGTTCACAGATGTGTGGCGGCTTCAATCCAGGA	2845
QY	5099	CGATCATCTGCGCAGGCTTTTCTTTCACTCGTGTGCTGTGTCGCGCGTATCTTTTGC	5158

Db	2844	CCCGCACTCTCGGAGGCCCTTCTGTTCCACACGCTGCTGTCGGCGGAAACCGTTTTC	2785
Qy	5159	AAAGCTTTGCATCTTAATACCTTAATTCAACGCACTGGAAGCGGAAATGGGGCTGTGGTTTTC	5218
Db	2784	GACCAAGCTTCATCTATCGCTGATCCATGCGCTGGAAACGCGCGGCGGCTCTGTGTTCCG	2725
Qy	5219	GGCGGGGGGATACCGGTCGCTGGTTATGGAATGAGCGCGGCTGTGTTTGCAATTTGGGCGG	5278
Db	2724	CAAGGGGGGACCAACCAAGCTGGTGGCGGGGATGTCCTGTTTGAGAGGCTTGGCGG	2665
Qy	5279	TGAATCTGTCYCAACCGCGAAAGTCACGCACTGGAAGCCGAGGGTAAACCGCATCAAGCGG	5338
Db	2664	CAGCTCTGCTGTAATGCCCGCGTCAAGCGGATCGAACAGAGGCGATGCGCCACAGGG	2605
Qy	5339	TGTCACACTGAAGGATGGGGCCGTTTGTGCGCCGCGCGTTCGTCGAATGCTACGCT	5398
Db	2604	CGTACACCTGTGACGAGCGGGCGATTGGCGGCGGATACGATGGCAGCAACGAGCAGCT	2545
Qy	5399	GGTGCATACCTACGATGCGCTGTATTAACCCAGCATCTCGCGGCGGTAAAGCGCGGCAAC	5458
Db	2544	GATGCACAGCTATGCGGCACTGTGCGGCGCATACCGCGGGGGGCGACCAAGGCGCGAT	2485
Qy	5459	GCTGAAGCGCAAGCGGATGAGCAACTGCGCTGTTGTACTCATTTTGTCTTATCATGC	5518
Db	2484	CCTGAACCGGAGCGGCTGCTGATGATGTGCTGTTCGTCGATTCGAGCTGTCCAAGCG	2425
Qy	5519	CCACCCGAGCTGGCGGACCAACAGATGTGTTGTGTCGCGGCTATGTAATTTGATCGA	5578
Db	2424	CCCCGAGAACCTGGCCCAACCAAGCTGATTTTGGCCCCGCGTCAAGAGGGCTGTGGA	2365
Qy	5579	TGAATCTTCAATGACAGCGAGCTGGCGGAAAGATTTCTGCTGTATCTGCATGCGCCCTG	5638
Db	2364	CGAATCTTCAACGGGCGCACGCTGGCGGACGATTTCTCATGTATCTGCAATTCGACCTCG	2305
Qy	5639	CTCCAGGATCCGTCGCTGGGACCGGGGGGCGTGGGAGATTTTAAAGTCGCGCGCGCTG	5698
Db	2304	CGTACCGGATCCAGCTCGGCCCCCGAGGGGATGTCCAGCATTAAGTCTCTGGGCCGT	2245
Qy	5699	GCCGCATCTCGTATCCGCGCGCATTTGACTGGCAACAGAAAGGGCGCGCTTGGCGCATCG	5758
Db	2244	TCCGCATCTGGCGCGCGCGCATGTGCATTTGGAAACCGAAGCCCCGGGCTATGCGGAGCG	2185
Qy	5759	CATTTTGTCTTATCTGAGAGACATATATGCCGGTCTGCGACAGCATTTAGTGACACA	5818
Db	2184	CATCTTCGAGAACTGAGCGCGCGGCATCTCCGACCTGGCAAGAACCTGACCGTCAAG	2125
Qy	5819	CCGTAATGTTTACGCGGTTTGAATTTTCGGAACAGCTCGACGCGCATAGGGCTCAGGCTT	5878
Db	2124	CCGCATCTTACGCCCGCCGATTTCAAGACCGAACTGTGGCCCATACGCGAGCGCTT	2065
Qy	5879	TTGCTCGAACCAATTTGACGCAAGCGCTGGTTCGGCGCGATACCGCATGCGCA	5938
Db	2064	CTCGATCGAGCGGATCTCGAAGCAATCGCTGTGTTCCGCCCATTAACCGCGACCGCGC	2005
Qy	5939	CATTACTPACTTTATCTGTGTGGGGGCTGAGACCGATCCCGGTGCGGTGTCGCAAGGCT	5998
Db	2004	GATCCCGAATCTTACATCTGTGGGGGGGGGACCGATCCGAGTCCGAGCATCCGGGGTGT	1945
Qy	5999	GATCGGCTCCCGCAAGGCGCGCCAGCTGATGATGGAGGATCTGACCGGATGAAACAA	6058
Db	1944	CGTTGGAGCGCCCAAGGCGCAGGGGCAAGTCACTGTGTCGGAACCTGGCGG---TCGATG	1888
Qy	6059	CCGCCGCTGATTTAGAGAGGTACCGCAACATGCGCGAGGGCTCAAAAGTTTCCGACG	6118
Db	1887	ACCGATCTGAAGGCGCATCTTCCGAAAGCGGCATTCGCGAGGGTGTGCAAAAGTTTGGCCAG	1828
Qy	6119	GCTACCCGGGCTATTGATCTTCAACGCGCGCGAGTACGCTGATGCTGTACGCTGTGT	6178
Db	1827	GGCGGCAAGCTGATGCCGCCCGCGCATCCGCAAGATACGGTCAAGCTCTGATGCTGATGCT	1768
Qy	6179	CGTCACTGTGACGATGTATGATGTGACGCTGGCGGAAGCGGCGACGACACGCG	6238

[illegible]





```

QY 5519 CCACCCGAGCTGGCGCAACCAACGCTGTGCTTGTGTCCTGCGGCTATGCTGAATTGATCGA 5578
Db 4850 CCCGGAACACCTGGACCCCAACACAGCTATCTTGCGCCCGGCTAACAGGGCGTGGAA 4791
QY 5579 TGAGATCTTCAATGAGCAAGCCAGCTGGCGGAGATTTTCGCTGATCTGATGCGCCCTG 5638
Db 4790 CGAATCTTCAAGCGGCGCAAGCCCTGCGGAGATTTTCGATGATCTGATGCGCCCTG 4731
QY 5639 CTCCAGCGATCCGTCCTGGCAACCGCGGCGCTGGCGAGTTTTCAGCTGCGCGCGT 5698
Db 4730 CGTACCGATCCAGCTGGCGCCCGGAGGGATGTCACAGCATTAACGTCCTTGCGCCGT 4671
QY 5699 GCGCATCTCGGTACCGCGCGCAATTGACTGGAACAGAGAGCGCGGCTTGGCGGATCG 5758
Db 4670 TCCCATCTGGGCGCGCGGATGATGATTGGAAAGCGAGGCGCGGCGCTATGCGAGCG 4611
QY 5759 CATTTTGTCTATCTGAGAGAGCACTATATGCGGGTCTGCGAGAGAGTTAGTAACA 5818
Db 4610 CATCTTGAAGAACTGAGAGCGCGCGCATCCCGACTGTCGCAAGCACTGACCGTCA 4551
QY 5819 CCGTATGTTTACGCGCTTGTATTTTCGAGACGCTGCAACGCGATCAGGGCTCAGCGTT 5878
Db 4550 CCGCATCTTACGCGCGCGGATTTACGACCGAATGTCGCGCCATCAACGCGACGCGCTT 4491
QY 5879 TTGCGTGAAGCCATTTTGAAGGAGCGCTGCTTCCGCGCGATTAACCGCGATGCGA 5938
Db 4490 CTCGATGAGCGGATCTGAGGCAATCGCTGCTTCCGCGGATTAACCGCGACGCGG 4431
QY 5939 CATTAACCTTATCTGCTGGGGGCTGAGACGATCCCGGTCGCGAGCGT 5998
Db 4430 GATCCGAACTTCAATCTGCTGGGGGCGGCAACGATCCGGGTGCGGCGGT 4371
QY 5999 GATCGGCTCCGGAAGGAGCGCGCGCTGATGATGAGGATCTGAACGGATGAACCA 6058
Db 4370 CGTTGGAGCGCCAGGCGCAAGGCGAGTATGCTGCGGACCTGGCGG---TCGCATG 4314
QY 6059 CCGCGCTGATTTGAGGAGGATCAGCGCAACCATGCGCGAGGGCTCCAAAGTTTGGCCAGC 6118
Db 4313 ACCGATCTGAGGGGACTTTCGAAAGCGCGCATCGCGAGGTTGCAAAAGCTTGGCGAG 4254
QY 6119 GCTACCGGCTATTTGATCTTCAACGCGCGCGAGTACGCTGATGCTGATGCTGCTGCT 6178
Db 4253 GCGGCAAGCTGATGCGCGCGCGCATCCGAGAGATACGATCATGCTTATGCTGCTGCTG 4194
QY 6179 CGTACTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6238
Db 4193 AGGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4140
QY 6239 GTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6298
Db 4139 GCGCGGAGCCAGCGCGCGCGCTGGGGCGCTGGCGCGGACAGCTGGCGCGCTGCAC 4080
QY 6299 GGTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6358
Db 4079 GAGAGCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4020
QY 6359 CTTCGCCAGAGCTGCTTTTGTATCTGATGATGATGATGATGATGATGATGATGATGAT 6418
Db 4019 TTCCCGGACTTTTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960
QY 6419 CGTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6478
Db 3959 GAATACCGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900
QY 6479 TTAATGATGAGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6538
Db 3899 GTGATGATGAGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3840
QY 6539 TTGGGCTGAGGCTTACGCTTACCAATATGACAGGAGATATGTTGAGAGCGCGAGAAAT 6598
Db 3899 CTGGGCTTGGGCTTACGCTTACCAATATGCTGCGAGGATGATGATGATGATGATGATGAT 3780
QY 6599 GCGCGTGTCTATCTGCGCAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGAT 6658

```

```

Db 3779 GGGCGCTGTCTATCTGCTGCTGCGGACTGCTGCGGAGCGG-----GGGAGCGTT 3729
QY 6659 GCGGATCCGCAACATGCGCGAGCGCTGGCCCGCTGGACGCGCTGTGTCGCGAGGCC 6718
Db 3728 GAGGATCCGCTGCTTGGAGCGCGCTTATTCGTCATCATCCGCGCTGTGAACGCGGCC 3669
QY 6719 GAGCGCTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6778
Db 3668 GAGCGCTATCATGCTTGGCGGCGGAGGCGCTTCCGATCTGCGCGCGCTGCGGCTGG 3609
QY 6779 GCGATGCGCACCGCGCGCGGCTTTACCGGGAATTGGCGTAAAGTGACATGCGGCT 6838
Db 3608 TCGATGCGCGCGCGCGCTGCTGCTATCTATCGGCAATCGGAGCGGATCCGCGAGGCG 3549
QY 6839 GCGCGGATGAGATACGCGCGAGCGCAACGATGAAGGCGAAAGCTGCGCTGCTGCTG 6898
Db 3548 CCGAGGCTATGCGCGAGGATCAGCATCGAAGGCTGCCAAGATCGGCGCTTTCGCGG 3489
QY 6899 AAAGTCCGCGCTGCGCTTACTTTCGCGCTTGTCTCATCCGAGGCGGCTTCCGCGGCT 6958
Db 3488 CGCGAGGCTTGAACGCGCGCGCATGCGCGCTGCGCGCGGCAAAATCAGCCGCGAGCG 3429
QY 6959 CTGTCGAGCGCTCGCG 6975
Db 3428 CTGTGACCCGACCGCG 3412

```

Search completed: December 8, 2004, 21:45:15  
 Job time : 34882 secs